INTELLIGENT MODEL FOR MONITORING, EVALUATING, AND RECOMMENDING STRATEGIES TO IMPROVE THE INNOVATION PROCESSES OF MSMES.

Submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Engineering

Ana GUTIERREZ

Universidad EAFIT

Escuela de Ingeniería

Medellín, Colombia

18/08/2023

Director of dissertation research:

Prof. Jose Aguilar Systems Engineering Department

Co-Directors of dissertation research:

Prof. Ana Ortega Marketing Department

Prof. Edwin Montoya Systems Engineering Department

Examining committee:

Prof. José Antonio Gutiérrez Universidad de Alcalá, Spain

Prof. Gabriela Marín Universidad de Costa Rica, Costa Rica

Prof. Cristina López González Universidad de Antioquia, Colombia

Dissertation committee chairperson:

Prof. Jorge Alonso Prieto Salazar Director of the Doctorate in Engineering

President of the Board of the Doctoral Program in Engineering:

Prof. Ricardo Taborda Rios Dean School of Engineering

Dissertation date: Febraury 29, 2024

Abstract

The research focuses on how to improve the innovation process in micro, small and medium-sized enterprises (MSMEs). The study is framed within the Smart Innovation paradigm. In this context, innovation is considered a relevant factor for organizational performance that allows the creation and improvement of competitive advantages through the implementation of new ideas, products, concepts, and services to increase market positioning. For organizations aiming to enhance innovation performance, using intelligent systems and artificial intelligence to guide the innovation process poses a challenge. To address this problem, the goal was to develop methodologies, models and approaches to support decision-making related to the intelligent management of the innovation process. To achieve this, specific objectives were defined. The first one is to design an intelligent model to support innovation processes in MSMEs. The second objective is to apply Artificial Intelligence (AI) techniques to customer data sources in social networks and organizational data of MSMEs, aiming to enhance the innovation process; The third objective is to develop an intelligent system to evaluate the innovation levels in MSMEs. The fourth objective is to instantiate a case study in the fashion cluster of the department of Norte de Santander and in the national context, as part of the applied methodology. To fulfill these objectives, research articles were developed. The process began with a literature review article on the current challenges in applying AI techniques to improve innovation processes in MSMEs. A proposed innovation model was made based on the different innovation models that exist in the literature, and the four research articles were written in compliance with the scientific standards that accredit them, to meet the specific objectives outlined in this doctoral thesis.

Each article evaluated the strategies/models using various data sets. The results demonstrated the capacity of the proposed methodologies and models for managing of innovation processes. For instance, the proposals enable the prediction of the level of innovation, and the definition of innovation problems, among other aspects, with positive results in performance metrics.

Keywords: Artificial intelligence, Machine learning, Innovation processes, Autonomic computing, Data analytics, Innovation decision-support system, Fuzzy cognitive maps.

Resumen

La investigación aborda el problema de cómo mejorar el proceso de innovación en las microempresas y las pequeñas y medianas empresas (PYME). El estudio del problema se sitúa en el marco del paradigma de la Innovación Inteligente. En este contexto, la Innovación es considerada un factor relevante para el desempeño organizacional que permite la creación y mejora de ventajas competitivas mediante la implementación de nuevas ideas, productos, conceptos, servicios para incrementar el posicionamiento en el mercado. Para las organizaciones que buscan mejorar el rendimiento de la innovación, el uso de sistemas inteligentes y la inteligencia artificial para guiar el proceso de innovación es un reto. Para abordar este problema, el objetivo era desarrollar metodologías, modelos y enfoques para apoyar la toma de decisiones relacionadas con la gestión inteligente del proceso de innovación. Para lograrlo, se fijaron objetivos específicos. El primero es diseñar un modelo inteligente para apoyar los procesos de innovación en las MIPYMES; el segundo es aplicar técnicas de Inteligencia Artificial (IA) a fuentes de datos de clientes en redes sociales y datos organizacionales de las MIPYMES, para mejorar el proceso de innovación; el tercero es desarrollar un sistema inteligente para evaluar el nivel de innovación en las MIPYMES; y el cuarto paso es instanciar un estudio de caso en el clúster de moda del departamento de Norte de Santander y en el contexto nacional, como parte de la metodología aplicada. Para ello, se desarrollaron artículos de investigación. Se inició con un artículo de revisión bibliográfica sobre los retos actuales en la aplicación de técnicas de IA para mejorar los procesos de innovación en las MIPYMES. Se realizó una propuesta de modelo de innovación basada en los diferentes modelos de innovación existentes en la literatura, y se redactaron los cuatro artículos de investigación cumpliendo con los estándares científicos que los acreditan, para cumplir con los objetivos específicos planteados en esta tesis doctoral.

En cada artículo se evaluaron las estrategias/modelos mediante el uso de diversos conjuntos de datos. Los resultados mostraron la capacidad de las metodologías y modelos propuestos para la gestión de los procesos de innovación. Por ejemplo, las propuestas permiten predecir el nivel de innovación y definir problemas de innovación, entre otras cosas, con buenos resultados en las métricas de rendimiento.

Palabras clave: Inteligencia artificial, Aprendizaje automático, Procesos de innovación, Computación autónoma, Análisis de datos, Sistema de apoyo a la toma de decisiones de innovación, Mapas cognitivos difusos

Scientific contributions

Several scientific articles were generated and published during the development of this research project.

Published articles:

1. Gutiérrez, A., Aguilar, J., Ortega, A. and Montoya, E, "Intelligent Systems in the Innovation Processes Of Textile MSMEs -A Systematic Literature Review," International Journal of Entrepreneurship, 2021, vol. 25, no. 5. https://www.abacademies.org/articles/intelligent-systems-in-the-innovation-processes-of-textile-msmesa-systematic-literature-review.pdf

   Q3 Scientific Journal Rankings

2. Gutiérrez, A., Aguilar, J., Ortega, A. and Montoya, E, "Autonomous cycles of data analysis tasks for innovation processes in MSMEs", Applied Computing and Informatics, 2022, Vol. ahead-of-print No. ahead-of-print. https://doi.org/10.1108/ACI-02-2022-0048

   Q1 Scientific Journal Rankings

Articles submitted to journals:

3. Gutiérrez, A., Aguilar, J., Ortega, A. and Montoya, E, "Sentiment analysis in the social networks for the definition of innovation problems in organizations", preprint submitted to Technology in Society.

   Q1 Scientific Journal Rankings

4. Gutiérrez, A., Aguilar, J., Ortega, A. and Montoya, E, "Using Fuzzy Cognitive Maps to evaluate the innovation in Micro, Small and Medium Enterprises", preprint submitted to Journal of Business Research.

   Q1 Scientific Journal Rankings

5. Gutiérrez, A., Aguilar, J., Ortega, A. and Montoya, E, “Definition of innovation problems in organizations using Data Analysis from Hybrid Sources: Social Networks and Organizational Databases”, preprint submitted to Technovation.

   Q1 Scientific Journal Rankings

Project context

This study was partially funded by the Colombian Administrative Department of Science, Technology and Innovation - COLCIENCIAS and Colombian Ministry of Science and Technology Bicentennial PhD Grant. The authors are grateful to physicians for using their knowledge and experience to build the models and interpreting the final results.

Acknowledgments

To God and the Virgin of Guadalupe for their multiple and countless blessings that help me grow as a professional and individual in body, mind, soul, and heart.

To my parents Luis Antonio and Ana Vicenta, and my siblings Alvaro and Lisbeth, along with my nephews Jesús David and Luciana who fill my life with happiness. Thank you for understanding my absence during the time I dedicated to my study; and for your generosity and patience.

To my true friends, you know who you are. Thank you for loving me so much, specially to my boss and friend Engineer Jorge Luis Orjuela for believing in me. Special thanks to my friends Merlyn, Lizet, Viviana, Francia; my grandmother for her constant prayers; and my aunt Blanca for their unconditional support when I needed it.

Thanks to Professor Eduard Puerto who believed in my talent and presented me with the opportunity to participate in this project. Thanks to him, I embarked on this beautiful doctoral research project. Thank you for making me a stronger and freer woman from today.

To my tutor José Lisandro Aguilar Castro, for his recommendations and corrections that contributed to complete this project. I am deeply grateful for his unconditional support and guidance throughout the academic process. His constant accompaniment facilitated my growth as a professional researcher in the field of Artificial Intelligence. His contributions were invaluable in this research.

To my co-tutors, Professor Ana María Ortega, thank you for your guidance, support in innovation and friendship. And to Professor Edwin Montoya, thank you for your support in this process.

To all of you, thank you for your support, teachings, and participation in this project. Thank you very much.

I extend my sincere thanks to my friends and fellow PhD students William, Claudia, Raúl, Rodrigo, and Jairo, as well as to Professors David Fernández and Pablo Muñoz from the University of Alcalá de Henares. Their warm hospitality during my doctoral stay in Spain was invaluable.

I would also like to express my gratitude to the esteemed faculties of Universidad EAFIT for their unwavering support and dedication, which has truly elevated the profession of Doctor of Engineering to one of which we can all feel immensely proud.

Contents

1 Introduction and research context 1

1.1 Problem statement and motivation ………………………………………………………………. 1

1.2 Research objectives ……………………………………………………………………………… 4

1.2.1 General objective ………...………………………………………………………………… 4

1.2.2 Specific objectives ………………………………………………………………………. 4

1.3 Contributions and research scope.……………………………………………………………. 4

1.4 Thesis organization ……………………………………………………………………………… 6

2 State of the art SLR Intelligent Systems in the Innovation Processes Of textile MSMEs – A Systematic Literature Review **8**

2.1 Motivation ……………………………………………………………..……………………… 8

2.2 Identification of the article ……………………………………………………………………. 9

2.3 Abstract ………………………………………………………………………………………. 9

2.4 Link to the full article ………………………………………………………………………… 9

3 An intelligent model based on Autonomous cycles of data analysis tasks for the management of innovation processes in MSMEs **10**

3.1 Motivation ……………………………………………………………………………………. 10

3.2 Identification of the article …………………………………………………………………… 11

3.3 Abstract ………………………………………………………………………………………. 11

3.4 Link to the full article ………………………………………………………………………… 12

4 Analysis of innovation problems in organizations using Artificial Intelligence (AI) techniques **13**

4.1 Motivation…………………………………………………………………………………...... 13

4.2 Sentiment analysis in the social networks for the definition of innovation problems in organizations ………………………………………………………………………………………. 14

4.2.1 Motivation ……………………………………………………………………………... 14

4.2.2 Identification of the article ………………………………………………………………. 14

4.2.3 Abstract ………………………………………………………………………………... 14

4.2.4 Link to the full article ……………………………………………………………………. 15

4.3 Analysis of innovation problems in organizations using Data Analysis from Hybrid Sources: Social Networks and Organizational Databases ……………………………..……………………. 15

4.3.1 Motivation ……………………………………………………………………………... 15

4.3.2 Identification of the article ………………………………………………………………. 16

4.3.3 Abstract. ……………………………………………………………………………… 16

4.3.4 Link to the full article ………..…………………………………………………………… 16

5 A Fuzzy Cognitive Map as an Intelligent Making Decision System to evaluate the innovation in MSMEs 17

5.1 Motivation ………..……………………………………………………………………… 17

5.2 Identification of the article …………………………...……………………………………. 17

5.3 Abstract ……………….……………………………………………………………………. 18

5.4 Link to the full article ………………………………………………………………………. 18

6 Conclusions 19

6.1 Summary …...…………………………………………………………………………….... 19

6.2 Limitations ……………. …………..……………………………………………………... 21

6.3 Future work ……………………………………………………………………………….. 22

References 24

Appendix A Intelligent Systems in the Innovation Processes Of textile MSMEs – A Systematic Literature Review 26

Appendix B Autonomous cycles of data analysis tasks for innovation processes in MSMEs **53**

Appendix C Sentiment analysis in the social networks for the definition of innovation problems in organizations **76**

Appendix D Definition of innovation problems in organizations using Data Analysis from Hybrid Sources: Social Networks and Organizational Databases **103**

Appendix E Using Fuzzy Cognitive Maps to evaluate the innovation in Micro, Small and Medium Enterprises **134**

Chapter 1

Introduction and research context

1. Problem statement and motivation

The increase in the speed of changes, both in the market and technology, makes it essential to consider innovation as a strategic variable, which is a condition for having a competitive advantage [7]. Although many enterprises recognize the importance of innovation, they do not implement it, either because they are very conservative or lack the necessary knowledge or support tools. Innovation has been extensively studied by the most successful enterprises worldwide, such as Apple, General Electric, Google, Honda, Microsoft, Facebook, Sony, etc., all of which have developed and applied different solution methods and strategies. For instance, leveraging artificial intelligence and the power of big data to improve growth and performance has enabled them to uncover effective innovation strategies.

One of the challenges facing Micro, Small & Medium Enterprises (MSMEs) today is applying innovation processes in all aspects. Knowing how to do it is just as important as implementing it into practice. Thus, innovation has become a strategic element allowing companies to enhance their competitive position. Its absence poses a serious problem in generating new products and processes [1]. In this sense, organizations must integrate innovative strategies into their operations. These strategies should not only focus on effectively managing innovation processes but also on enhancing adaptability and proactively anticipating disruptions. Such actions enable organizations to constantly improve their competitive advantages promptly [2].

MSMEs are one of the business sectors that require continual innovation to improve business results [3]. On the other hand, organizations fostering a culture of innovation experience a significant 32% increase in productivity. They also substantially mitigate turnover risks by 14%, potentially multiplying their annual revenue by fivefold. These findings were highlighted in the 2019 "Cultures of Innovation Model" study conducted by the renowned consulting firm, Great Place to Work [19]. Today, enterprises need to innovate to survive. If enterprises do not apply innovation to their processes and products, they will be in danger of interrupting their business operations.

Then again, Colombia's investment level in innovation activities is still low, although it is the fifth most innovative economy in Latin America among 18 countries [4]. Despite this ranking, it allocates less than 0.8 percent of GDP (Gross Domestic Product) to this area, with only 35 percent of those resources come from the private sector, according to the Global Innovation Index, published in July 2019 and carried out annually by Cornell University (United States) and the World Intellectual Property Organization (WIPO) [5].

Department of Norte de Santander currently has various MSMEs established in the fields of Agroindustry, Commerce, Tourism, Health, TIC, and Fashion, among others. To take advantage of the competitive advantages of the region, and considering Michel Porter's Theory in his book "On Competition" (page 199), where he says, "A cluster is a concentration of enterprises and institutions in the same business, with multiple relationships between them and with other agents in the industry, united by common features and complementarities", in Cúcuta city, North of Santander, the northern fashion cluster has been created to develop the competitiveness of the enterprises in the fashion system of North of Santander. The competitive scope of the cluster lies within the fashion industry, offering favorable profit margins. This sector demands the integration of trend-responsive design, streamlined logistics, efficient production processes, and effective retail management, as key value-added components. Another main objective is to accelerate the innovation processes, and therefore, competitiveness across, local, regional, and national levels. Enterprises must devise new strategies aimed at enhancing competitiveness, by improving productivity and delivering products with higher added value. This cluster has around 1,671 enterprises, including manufacturers of flat fabric (jeans and shirts) and knitted fabric (cotton and lycra). The cluster comprises various stakeholders, including women's footwear manufacturers, leather goods suppliers, marketing professionals, unions, and support entities like universities and chambers of commerce. Out of a total of 1,671 companies, 615 are engaged in the production of both footwear and clothing, collectively providing employment to approximately 3,500 individuals. Additionally, it is estimated that the marketing of these products creates 2,000 more jobs [6].

On the other hand, at the regional level, the MSMEs established in the Agro-industrial, Commerce, Tourism, Health, ICT, and Fashion fields of the Department of North of Santander have a moderately acceptable level of Information and Communication Technology, according to the Development Plan for Norte de Santander 2016-2019 “A Productive North for All”. The MSMEs have conventional information systems, internet access, and use of social networks, but it is necessary to investigate how to exploit these resources so that they can develop strategic policies that allow them to innovate in their processes and products. It is essential to develop a strategy to generate added value to the product offerings of small and medium-sized national and regional companies, so that differentiation is the basic element of competitiveness. For this reason, this project proposes to develop intelligent models that help reduce the innovation gap of companies through data analysis. Intelligent models that can predict the level of innovation, define innovation problems using different data sources, among other things.

In short, there is a need to develop different knowledge models to improve the innovation processes. The expectation is to develop knowledge models that can be integrated into an architecture with autonomous tasks in each of the sub-processes of an innovation process [8].

1.2 Research objectives

The general objective and the specific objectives of this research are presented below.

1.2.1 General objective

Create an intelligent system for monitoring, evaluating, and defining of strategies to improve the innovation processes of MSMEs in the North of Santander Department based on data analysis about MSMEs and their clients.

1.2.2 Specific objectives

1. Design an intelligent model to support innovation processes in MSMEs.

2. Apply AI techniques to customer data sources on social networks and organizational data of MSMEs to improve innovation processes.

3. Develop an Intelligent System to evaluate the innovation level in MSMEs.

4. Instantiate the model through a case study in the fashion cluster of the North of Santander Department and in the national context.

1.3 Contributions and research scope

This research focuses on the development of models, methodologies, and computational approaches to support the decision making in the innovation processes (the case study is MSMEs in the textile industry). This study makes several contributions presented in this section. First, a systematic literature review (SLR) was conducted to identify the challenges and opportunities associated with improving the innovation processes [9]. The review focused on two areas of interest. AI, and Innovation models, which are relevant to the management and improvement of innovation processes. With this in mind, challenges and opportunities for future work are identified. Based on the information reported in the SLR, specific objectives related to the architecture of an intelligent model based on autonomous cycles of data analysis tasks for the management of innovation processes are proposed.

In this sense, the use of AI, has the potential to improve the innovation process. AI can discover patterns, synthesize information, draw conclusions, make predictions, or perform problem-solving tasks, among other things [8]. The integration of AI in the innovation processes allows for improving its first sub-process, problem definition, to make better decisions based on customers' opinions, comments, and needs. AI techniques were applied to customer data sources on social networks and organizational data of MSMEs to enhance innovation processes. The datasets are related to social network comments, customer satisfaction surveys, and PCCS (petitions, complaints, claims, suggestions), and this information is used to define innovation problems using the 5W methodology and the Natural Language Processing paradigm. In addition, a tool was developed to assess the innovative capacity of organizations based on FCM, particularly, in MSMEs.

In general, the data sources used to develop these knowledge models were obtained at the regional level from the company RAMARA JEANS, and at the national level from the company OFFCORSS, to establish a case study in the fashion cluster of the North of Santander Department and in the national context. For both case studies, the data for the last five years (from 2018 until 2022) were used. Other data sources used were from the Departmental Innovation Index for Colombia, the World Innovation Index, and the MSMEs based in the Valle de Aburrá, specifically in Ruta N [11], to develop an Intelligent System to evaluate the level of innovation in MSMEs. The models and systems were evaluated using particular scenarios determined by innovation experts.

For this purpose, several research articles were developed to fulfill the objectives proposed in this thesis. Initially, a literature review article was conducted on the current challenges in the application of AI techniques to improve innovation processes in MSMEs. Additionally, a proposal was made for an innovation model based on the different innovation models existing in the literature. On the other hand, for the first specific objective, an article was written to specify the innovation process using the concept of

Autonomous Cycle of Data Analysis Tasks (ACODATs) to manage it. In this article, a multidimensional data model is defined for the management of an innovation process, which stores the necessary information of the organization and the context. To meet the second specific objective, two articles oriented to the analysis of innovation problems in organizations using Artificial Intelligence (AI) techniques are proposed. The first one uses sentiment analysis in social networks for the definition of innovation problems in organizations and the second one analyzes innovation problems in organizations by performing data analysis from hybrid sources: social networks and organizational databases. For the third objective, a Fuzzy Cognitive Map (FCM) was developed as an Intelligent System to evaluate the level of innovation in MSMEs. Each article evaluated, the strategies/models using various data sets.

Finally, all contributions made in this research are represented in several research articles. A total of five (5) scientific articles were generated, of which two (2) are published and the other three (3) are under review.

1.4 Thesis organization

This thesis is presented as a collection of articles developed to meet each of the proposed objectives. Chapter 2 describes the results of our SLR. Chapter 3, Chapter 4, and Chapter 5 correspond to the fulfillment of the first, second, third and fourth objectives, respectively. One article was generated for the first sub-objective, two articles for the second sub-objective, and one article for the third sub-objective. These articles will be presented in each chapter.

A brief description of each chapter is presented below. Chapter 2 describes the results of our SLR on AI and sensing techniques for the innovation processes. This SLR allowed us to identify trends, challenges, and research opportunities in this field. Chapter 3 shows the architecture of an intelligent model based on ACODAT for the management of innovation processes in MSMEs to achieve the first and fourth specific objectives. Chapter 4 presents two articles to meet the second and fourth specific objectives proposed in this thesis. The first article corresponds to sentiment analysis in social networks for the definition of innovation problems in organizations. The second article corresponds to the analysis of innovation problems in organizations using Data Analysis from hybrid sources: social networks and organizational databases. Chapter 5 presents one article about the FCM's utilization as an Intelligent System to evaluate the innovation level in MSMEs.

Finally, Chapter 6 presents a summary of all the articles' conclusions presented in the previous sections. It also shows the limitations of our research and possible future works.

Chapter 2

State of the art on Intelligent Systems for the Innovation Processes of Textile MSMEs

2.1 Motivation

This chapter presents the SLR proposed for identifying and evaluating the available research about intelligent systems used in MSMEs, aiming to understand knowing the trends, current works, and publications in that field. For that, the paper proposed three research questions [9]: 1. How can MSMEs innovation processes be improved? 2.What intelligent systems have been developed that contribute to improving the innovation of MSMEs? 3.How has AI been used in the textile industry?

Results of the previous studies demonstrated the importance of knowledge management within organizations, and its impact on the innovation processes. Consequently, it is necessary to create tools for optimal knowledge management; particularly, the use of analytical tools to support decision-making during the innovation processes is a big challenge, both in theoretical and practical terms. There were two contributions to this work: the determination of the gaps that exist in the field of research on innovation in organizations, and the identification of how innovation processes can be improved by applying AI. A proposal for an innovation model was also generated based on the different innovation models proposed in the literature. The main conclusion was that it´s possible to implement systems that allow companies to improve their innovation processes by applying data analysis tools.

2.2 Identification of the article

Gutiérrez, A., Aguilar, J., Ortega, A. and Montoya, E, "Intelligent Systems in the Innovation Processes of Textile MSMEs - A Systematic Literature Review," International Journal of Entrepreneurship, 2021, vol. 25, no. 5.

2.3 Abstract

Innovation has been considered a relevant factor for organizational performance that allows the creation and improvement of organizational competitive advantages, through the implementation of new ideas, products, concepts, services, and practices to increase productivity and market value. For organizations in the pursuit of innovation performance, the utilization of intelligent systems and artificial intelligence to guide the innovation process is a challenge. This presents a systematic literature review that analyses how intelligent systems have been used to improve the innovation capacities in Micro, Small & Medium Enterprises. The results show that even though there are relevant papers on the subject, there is not an established body of knowledge sufficient to ensure clear guidelines to improve the innovation process and performance based on intelligent systems. This paper exposes and discusses several challenges in this domain.

2.4 Link to the full article

See Appendix A and https://www.abacademies.org/articles/intelligent-systems-in-the-innovation-processes-of-textile-msmesa-systematic-literature-review.pdf

Chapter 3

An Intelligent Model based on Autonomous Cycles of Data Analysis Tasks for the Management of Innovation Processes in MSMEs.

3.1 Motivation

Given the importance of innovation in MSMEs, and the existing opportunities to exploit data from the organizations, contexts, and strategies based on data could be defined to build data-driven models to guide the innovation processes. One of these strategies is the utilization of the concept of "Autonomous Cycles of Data Analysis Tasks" (ACODATs) [12][13][14], which allows generating knowledge models convenient for the management of innovation processes using different data sources. This article proposed the concept of "Autonomic Cycle for innovation Processes", which defines a set of tasks of data analysis whose objective is to improve the innovation processes in MSMEs. Notably, the ACODAT was instanced to determine innovation problems in the textile industry [8]. Thus, this chapter presents the paper for the fulfillment of the first objective, and for the partial fulfillment of the fourth objective because there is a case study in the fashion cluster of the North of Santander Department.

3.2 Identification of the article

Gutiérrez, A., Aguilar, J., Ortega, A. and Montoya, E, "Autonomous cycles of data analysis tasks for innovation processes in MSMEs", Applied Computing and Informatics, 2022, Vol. ahead-of-print No. ahead-of-print. https://doi.org/10.1108/ACI-02-2022-0048

3.3 Abstract

Purpose – The authors propose the concept of “Autonomic Cycle for innovation processes,” which defines a set of tasks of data analysis, whose objective is to improve the innovation process in micro-, small and medium- sized enterprises (MSMEs).

Design/methodology/approach – The authors design autonomic cycles where each data analysis task interacts with each other and has different roles: some of them must observe the innovation process, others must analyze and interpret what happens in it, and finally, others make decisions in order to improve the innovation process.

Findings – In this article, the authors identify three innovation sub-processes which can be applied to autonomic cycles, which allow interoperating the actors of innovation processes (data, people, things and services). These autonomic cycles define an innovation problem, specify innovation requirements, and finally, evaluate the results of the innovation process, respectively. Finally, the authors instance/apply the autonomic cycle of data analysis tasks to determine the innovation problem in the textile industry.

Research limitations/implications – It is necessary to implement all autonomous cycles of data analysis tasks (ACODATs) in a real scenario to verify their functionalities. Also, it is important to determine the most important knowledge models required in the ACODAT for the definition of the innovation problem. Once determined this, it is necessary to define the relevant everything mining techniques required for their implementations, such as service and process mining tasks.

Practical implications – ACODAT for the definition of the innovation problem is essential in a process innovation because it allows the organization to identify opportunities for improvement.

Originality/value – The main contributions of this work are: For an innovation process is specified its ACODATs in order to manage it. A multidimensional data model for the management of an innovation

process is defined, which stores the required information of the organization and of the context. The ACODAT for the definition of the innovation problem is detailed and instanced in the textile industry. The Artificial Intelligence (AI) techniques required for the ACODAT for the innovation problem definition are specified, in order to obtain the knowledge models (prediction and diagnosis) for the management of the innovation process for MSMEs of the textile industry.

3.4 Link to the full article

See Appendix B and https://www.emerald.com/insight/content/doi/10.1108/ACI-02-2022-0048/full/html

Chapter 4

Analysis of Innovation Problems in Organizations using Artificial Intelligence (AI) Techniques

4.1 Motivation

Innovation serves as a fundamental pillar for the continuous growth and competitiveness of organizations spanning diverse industries. Nevertheless, innovation processes come with their own inherent challenges and intricacies. Notably, the incorporation of AI techniques presents a compelling solution to enhance these innovation processes significantly. For example, AI can identify patterns, synthesize information, draw conclusions, make predictions, and engage in problem-solving tasks, among other things [15]. Moreover, the substantial data generated in the digital era both poses a challenge and offers an opportunity. Organizations must adopt effective data management and analytics strategies to harness the full potential of this valuable resource while mitigating associated risks. This chapter presents two articles designed to address the second and fourth specific objectives. The first article presents a sentiment analysis of social networks to define organizations' innovation problems. The second article presents an analysis of innovation problems in organizations using data analysis from hybrid sources: social networks and organizational databases.

4.2 Sentiment Analysis in the Social Networks for the Definition of Innovation Problems in Organizations

4.2.1 Motivation

Sentiment analysis on social networks allows for the evaluation of the opinions disseminated through this media. Social networks facilitate the communication efforts of firms to transmit their values, attract new consumers, get feedback from discontent customers, or provide selling info [16]. On the other hand, the definition of the problem determines the scope of an innovation process. Therefore, an adequate and comprehensive definition of an innovation problem is a key phase to be considered in any kind of innovation process [17][18]. According to the above, it is possible to affirm that social networks are useful for a company to have extra sources of information, which will allow identifying and defining problems of innovation. This article used sentiment analysis on social networks for the definition of innovation problems in organizations. Thus, in this section, the first article is presented to fulfill the second and fourth specific objectives, which considers a national case study.

4.2.2 Identification of the article

Gutiérrez, A., Aguilar, J., Ortega, A. and Montoya, E, "Sentiment analysis in the social networks for the definition of innovation problems in organizations", preprint submitted to Technology in Society.

4.2.3 Abstract

In recent years, on social networks, people can express their thoughts, write their emotions, and generate content about many topics. In order to discover the value of this information, in this article, we propose a sentiment analysis system for the social networks based on the "autonomic cycle of data analysis tasks" paradigm, to define innovation problems in an organization. This autonomic cycle has a set of tasks used for gathering and handling large unstructured data from social media, for the definition of innovation problems through sentiment analysis. The tasks of the autonomic cycle filter the negative tweets and determine their keywords. Then, they calculate the groups/clusters in the negative tweets, and for each cluster, they analyze each centroid to obtain additional information to answer the 5 questions of the W model (What, Where, When, Why, Who) that describe a problem. Particularly, considering this information on the clusters, the last task defines customer-oriented innovation problems. Finally, this article analyzes a case about tweets in a fashion enterprise, with results very encouraging.

4.2.4 Link to the full article

See Appendix C.

4.3 Analysis of Innovation Problems in Organizations using Data Analysis from Hybrid Sources: Social Networks and Organizational Databases

4.3.1 Motivation

The motivation for exploring data analytics from hybrid sources, particularly a blend of sources such as social networks and organizational databases, stems from the necessity of adopting a comprehensive, data-centric approach to defining innovation challenges. This multifaceted analysis empowers organizations to unearth concealed patterns, identify influential ideas, among other things, to make well-founded decisions that fuel the innovation process. As innovation continues to gain prominence in the business arena, harnessing the potential of hybrid data sources emerges as a strategic imperative for organizations aspiring to achieve lasting growth and prosperity. This paper presents an ACODAT for defining innovation problems based on sarcasm analysis in satisfaction rating surveys, and customer comments in social networks and PCCS systems. Thus, in this section, the second article is presented to fulfill the second and fourth specific objectives.

4.3.2 Identification of the article

Gutiérrez, A., Aguilar, J., Ortega, A. and Montoya, E, "Definition of innovation problems in organizations using Data Analysis from Hybrid Sources: Social Networks and Organizational Databases", preprint submitted to Technovation.

4.3.3 Abstract

Nowadays, we can express the experience we have lived with the products we use. Most of the time, we interact with brands and let our likes and dislikes be seen on digital platforms, either by interacting with social networks, filling out satisfaction surveys, or registering requests, complaints, and claims. To get the value of all available customer data of the companies, in this article, we propose to study the data from different sources using an autonomic cycle of data analysis tasks to define innovation problems in an organization. The tasks of the autonomic cycle filter the customer comments from different sources (e.g., from social networks, PCCS (petitions, complaints, claims, suggestions) systems of organizations, etc.), obtain their keywords, and analyze the patterns of the users to answer the questions of the 5W methodology (what, who, where, when and why?), in order to define innovation problems. Finally, this article analyzes a case study of a fashion company using its PCCS system and comments of Twitter, to identify useful information. Part of the information discovered was the reasons for customer returns, merchandise delivery problems, shipment failures, failure to respond timely to customers, among other things. With this information, the autonomic cycle is able to define customer and organization-oriented innovation problems, in order to respond to these identified problems.

4.3.4 Link to the full article

See Appendix D.

Chapter 5

A Fuzzy Cognitive Map as an Intelligent Making Decision System to Evaluate the Innovation in MSMEs.

5.1 Motivation

The reason behind creating an innovation evaluation model is to provide a valuable tool that aids managers, entrepreneurs, and government leaders at various levels. This model serves to discern and quantify innovation levels within organizations and countries on a global scale, offering crucial insights not only for these decision-makers in organizations, but also, for specific national departments tasked with fostering and regulating innovation. By doing so, it enables us to define informed strategies and policies aimed at driving innovation and fostering economic growth. By analyzing data such as innovative activities, basic research, and external context, the evaluation model can identify three levels of innovation: low, medium, and high. Thus, in this section, the article is presented to fulfill the third specific objective.

5.2 Identification of the article

Gutiérrez, A., Aguilar, J., Ortega, A. and Montoya, E, "Using Fuzzy Cognitive Maps to evaluate the innovation in Micro, Small and Medium Enterprises", preprint submitted to Journal of Business Research.

5.3 Abstract

The field of innovation covers not only the development of new products, but also new internal processes, the management of new market positioning initiatives, and the generation of new concepts and platforms, sometimes based on new paradigms. In the innovation process, organizations should measure activities and evaluate innovation performance. This will allow them to continuously improve their processes, methods, and capabilities, and increase product and process innovation in the future. Controlling the evaluation of innovation using intelligent systems can be challenging for companies seeking innovation performance. This paper presents a fuzzy cognitive map for improving the assessment of innovation in SMEs. The previous results in the literature show that although there are relevant advances on this topic, there is not enough knowledge to provide clear guidelines for evaluating innovation and improving performance in an organization using intelligent systems. However, we get very good results with our approach based on fuzzy cognitive maps in three case studies to determine the level of innovation in an organization. Additionally, this article reveals and addresses several challenges in this field.

5.4 Link to the full article

See Appendix E.

Chapter 6

Conclusions

This chapter presents a summary of the results of all the work presented above. In addition, it shows limitations and research opportunities for the future.

6.1 Summary

The integration of information technologies, particularly AI, in the field of innovation, represents a critical step for organizations. It constitutes a vital milestone for organizations aiming to flourish in the swiftly evolving business environment, ensuring their market presence and the attainment of a competitive edge. The most appropriate AI technologies for MSMEs to support and improve innovation processes with the use of data analysis tools were explored and defined. In this sense, the following findings were obtained:

First, it was possible to propose the automation of the innovation processes in MSMEs through the definition of ACODATs, which used different data sources to build knowledge models about the innovation processes (e.g., predictive and descriptive models). It was possible to generate knowledge for the organization, not only to identify a problem but also, to identify where it happened, when it happened, and the impact it has on the organization. Furthermore, the use of ACODAT allows for more informed, data-driven decisions.

In addition, it scales to handle larger volumes of data and adapts to evolving data analysis requirements as MSMEs grow. Thus, through continuous data analysis, ACODATs help MSMEs adapt to changing market conditions and customer preferences, ensuring long-term sustainability. Undoubtedly, the definition of an ACODAT for the identification of innovation problems was a great contribution for companies.

Secondly, a protocol for the application of the ACODAT was developed, and a real case was addressed based on data obtained from a social network and organizational databases of a Colombian company dedicated to the manufacture, marketing, and sale of clothing. There, each of the ACODAT tasks was implemented to define innovation problems, to verify its functionality and to evaluate the quality of the results. This has empowered the company to make informed decisions on how to effectively address these problems.

Hence, by automating data analysis tasks and focusing on relevant data sources, the company can allocate its resources more efficiently, directing efforts toward areas that require attention. ACODAT is a powerful tool to harness the value of information from social networks, organization, surveys, etc., in order to define innovation problems with precision, and drive strategic decision-making.

Finally, an FCM was proposed for the evaluation of the level of innovation in an organization, which can be applied in different contexts, as evidenced by the Colombian case studies and the global case study, with very good results. It is important to highlight that this is the first work that proposes an FCM for the evaluation of innovation using different types of data as the innovative activities, and external contexts, among others, as the macro variables of the organizational environment that influence the development of innovative competencies, based on the Oslo Manual.

The objective of the research aimed at "creating an intelligent system for monitoring, evaluating and defining strategies to improve the innovation processes of MSMEs in the Department of Norte de Santander based on the analysis of data on MSMEs and their customers", was developed through the design of an intelligent model applying AI techniques and implementation of an intelligent system to evaluate innovation in MSMEs. It achieved an accuracy of 80% in Colombian cases and 92% % in global cases to assess the levels of innovation in organizations. This implies that it can be applied in any industry context.

In conclusion, there are several impacts of an intelligent system such as FCM for innovation assessment: it empowers organizations to make data-driven decisions, fosters innovation and improves their competitive position in the global landscape. Its high accuracy, adaptability and alignment with international standards make it a valuable asset for organizations striving for innovation excellence.

6.2 Limitations

One of the main limitations in the development of this work was that the databases used were at the local level, relying solely on the information provided by the studied organization, which consequently resulted in the utilization of static data within a defined date range. This limitation can hinder the ability to capture real-time or evolving trends in innovation, which are essential for making proactive decisions.

Another limitation is that some of the datasets were incomplete. For example, not all customers express their opinions online, which could lead to an incomplete representation of customer sentiments. This work used data from different sources such as social networks, and PCCS systems. This limitation is especially relevant when conducting sentiment analysis, as it may not capture the sentiments of a significant portion of the customer base. Incomplete datasets may limit the depth of insights that can be derived from the analysis. Understanding the full spectrum of customer sentiments and experiences is essential for making well-informed decisions and addressing potential issues.

Another limitation might be the potential biases or incomplete representation of customer opinions if certain demographics or platforms are not well-covered. Finally, data is constantly evolving. Thus, the clustering models need to adapt to new information, features, and trends to maintain their effectiveness over time. Models that can continuously learn and incorporate new information are better equipped to provide accurate and actionable insights over time, making them valuable assets for data-driven decision-making in dynamic environments.

6.3 Future Work

It is necessary to implement our ACODATs in other real scenarios to verify their functionalities. For example, Customer Service and Support could be applied to analyze customer interactions, such as emails, chat logs, and call recordings, to identify common issues, customer sentiment, and areas for service improvement. Another area can be HealthCare to analyze patient records, medical literature, and social media data to identify healthcare trends, patient sentiment, and areas where medical research is needed. Thus, the versatility of ACODATs makes them a valuable tool for data analysis and decision-making in a wide range of real-world scenarios. Also, it is important to define the ACODAT for the other sub-processes of our general innovation model. Once determined, it is necessary to define the relevant mining techniques required for their implementations. This involves the collection and analysis of data, ensuring data privacy and security in compliance with regulations.

Furthermore, the use of other sentiment analysis frameworks, for example, those based on Sentic computing, will allow for increasing the precision of sentiment analysis systems. For example, the integration of Sentic computing can lead to a better grasp of how sentiment is influenced by the surrounding context.

Future work must explore the integration of additional information from the organization, for example, operational data provides real-time information about the day-to-day activities of the organization. It can include data on production, inventory, supply chain, and more. This data helps in monitoring and optimizing processes for innovation, identifying bottlenecks, and streamlining operations. On the other hand, financial records and sales data provide information about customer behavior and product performance, and market demand. It can be used to identify product trends, customer preferences, and potential opportunities for product or service innovation. Also, beyond internal financial records, external financial data, such as market trends, industry benchmarks, and competitive information, can help in assessing the competitive landscape and identifying market gaps or emerging opportunities for innovation.

Other future work is to implement an online tool with an intuitive, user-friendly and visually appealing user interface that can perform the entire process of problem definition through these various data analysis tasks.

This tool will encourage cross-functional collaboration and participation in problem-solving and innovation initiatives. Second, it will simplify the process of importing, cleaning, and processing data from various sources, and will save time and reduce the technical barriers associated with data analysis. Finally, the tool will provide real-time updates on data trends and insights, enabling timely decision-making and responsiveness to change circumstances.

Finally, another future work is to implement an automatic check for innovation problems that checks whether the formulation is appropriate and correct for the context in which it is used. For example, implement semantic analysis to ensure that the problem formulation aligns with the context and goals of the organization. This will involve checking for logical inconsistencies to ensure that the problem is well-defined, using formal rules and constraints.

References

[1] Nuchera, H., “La gestión de la innovación como proceso”, U. de Rica (Ed.), Conocimiento, Innovación y Desarrollo, Impresión Gráfica del Este, Costa Rica, 2011, Vol. 1, no. 99-122

[2] Robayo, P., “La innovación como proceso y su gestión en la organización: una aplicación para el sector gráfico colombiano". Suma de Negocios, 2017, Vol. 7, no. 125-140

[3] Weldeslasie, D., Ahmed, M., “The application of expert systems in manufacturing sector", International Journal of Innovative Technology and Exploring Engineering, 2019, Vol 9, no. 4744-4749.

[4] “Estudios de la OCDE de las Políticas de Innovación: Colombia”. OCDE 2020

[5] Antosova, G., Cifuentes, J., “Propiedad Intelectual y Procesos de Innovación en el marco del Objetivo 17 de Desarrollo Sostenible”. Civilizar Ciencias Sociales y Humanas, 2019, Vol. no. 117-130. https://doi.org/10.22518/usergioa/jour/ccsh/2019.2/a10

[6] Innpulsa, Comisión Regional de Competitividad de Norte de Santander, Cámara de Comercio de Cúcuta, Fundación de Estudios Superiores Comfanorte, Ministerio Industria y Comercio, 2014, Iniciativa el Norte de la Moda.

[7] Kaltoff, O., Nonaka, T., Nueno, P., “La luz y la sombra: La innovación en la empresa y sus formas de Gestión”. Ediciones Deusto, 1998, Bilbao.

[8] Gutiérrez, A., Aguilar, J., Ortega, A. and Montoya, E, "Autonomous cycles of data analysis tasks for innovation processes in MSMEs", Applied Computing and Informatics, 2022, Vol. ahead-of-print No. ahead-of-print. https://doi.org/10.1108/ACI-02-2022-0048

[9] Gutiérrez, A., Aguilar, J., Ortega, A. and Montoya, E, “Intelligent Systems in the Innovation Processes Of Textile MSMEs - A Systematic Literature Review,” International Journal of Entrepreneurship, 2021, vol. 25, no. 5.

[11] Ruta N: innovation and business center of Medellín “Raw data on Innovation in Valle de Aburrá organizations”, 2020, https://www.rutanmedellin.org/

[12] Sánchez, M., Aguilar, J., Cordero, J., and Valdiviezo, P. “Basic Features of a Reflective Middleware for Intelligent Learning Environment in the Cloud (IECL)”, in proceeding Asia-Pacific Conference on Computer Aided System Engineering (APCASE), 2015.

[13] Sánchez, M., Aguilar, J., Cordero, J., Valdiviezo-Díaz, P., Barba-Guamán, L., Chamba-Eras, L., “Cloud Computing in Smart Educational Environments: Application in Learning Analytics as Service”. In: Rocha, Á., Correia, A., Adeli, H., Reis, L., Mendonça Teixeira, M. (eds) New Advances in Information Systems and Technologies. Advances in Intelligent Systems and Computing, 2016, vol. 444, no. 993-1002.

[14] Aguilar, J., Cordero, J., Buendía, O., “Specification of the autonomic cycles of learning analytic tasks for a smart classroom”. Journal of Educational Computing Research, 2018, Vol. 56, no. 866-891.

[15] Füller, J., Hutter, K., Wahl, J., Bilgram, V., and Tekic, Z., “How AI revolutionizes innovation management – Perceptions and implementation preferences of AI-based innovators,” Technological Forecasting and Social Change, 2022, vol. 178, doi: 10.1016/j.techfore.2022.121598.

[16] Paroubek, P., and Pak, A., “Twitter as a Corpus for Sentiment Analysis and Opinion Mining”, In Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10), Valletta, Malta. European Language Resources Association (ELRA), 2010.

[17] Wilkerson, B., and Trellevik, L., “Sustainability-oriented innovation: Improving problem definition through combined design thinking and systems mapping approaches,” Thinking Skills and Creativity, 2021, Vol. 42, https://doi.org/10.1016/j.tsc.2021.100932.

[18] Hoolohan, C., Browne, A., “Design thinking for practice-based intervention: Co-producing the change points toolkit to unlock (un)sustainable practices,” Design Studies, 2020, Vol. 67, no. 102–132, https://doi.org/10.1016/j.destud.2019.12.002.

[19] Great Place to Work. (2019). Culturas de Innovación: Informe global de investigación 2019.

Appendix A

Intelligent Systems in the Innovation Processes Of textile MSMEs – A Systematic Literature Review

Volume 25, Special Issue **Print ISSN: 1099-9264**
Online ISSN: 1939-4675

INTELLIGENT SYSTEMS IN THE INNOVATION PROCESSES OF TEXTILE MSMES–A SYSTEMATIC LITERATURE REVIEW

Ana Gutiérrez, Universidad EAFIT
Jose Aguilar, Universidad de Los Andes, Universidad de Alcalá
Ana Ortega, Universidad EAFIT
Edwin Montoya, Universidad EAFIT

ABSTRACT

Innovation has been considered a relevant factor for organizational performance that allows the creation and improvement of organizational competitive advantages, through the implementation of new ideas, products, concepts, services, and practices to increase productivity and market value. For organizations in the pursuit of innovation performance, the utilization of intelligent systems and artificial intelligence to guide the innovation process is a challenge. This presents a systematic literature review that analyses how intelligent systems have been used to improve the innovation capacities in Micro, Small & Medium Enterprises. The results show that even though there are relevant papers on the subject, there is not an established body of knowledge sufficient to ensure clear guidelines to improve the innovation process and performance based on intelligent systems. This paper exposes and discusses several challenges in this domain.

Keywords: Innovation Processes, Smart Innovation Capabilities, Intelligent Systems, Artificial Intelligence, Micro-Small and Medium Enterprises, Textile Industry

INTRODUCTION

Currently, it is difficult to think about moving forward and staying in the market without innovating. Regardless of the type of company, organizational advancement and development might be linked to its innovation activities and results (OECD, 2018). Micros, Small & Medium Enterprises (MSMEs), are some of the business sectors that need permanent innovation the most, to improve business results (Weldeslasie & Ahmed, 2019).

In general, innovation is considered a critical factor for companies´ survival and success due to the necessity to adapt faster and better to the changing needs of the environment. The reality is that MSMEs, due to their particular characteristics, in many cases obviate the implementation of a comprehensive plan of assets linked to innovation, technology management, and investment in research and development (Ferrer et al., 2015). Currently, the importance of innovation is increasingly evident as a consequence of rapid changes, which requires managers to explore and implement actions that allow them to tackle market´s needs, by designing and implementing adequate strategies in their companies (Sánchez, 2016). Organizational managers must recognize the importance of innovation and commit with the necessary internal changes, related to human capital development and creativity, at individual and group levels, being this a key factor for success (Flórez, 2019).

The innovation has been widely studied by the most successful companies worldwide, such as Apple, General Electric, Google, Honda, Microsoft, Facebook, and Sony, which have developed and applied different methods and solution strategies. For example, based on Artificial Intelligence (AI), some of these companies have managed to boost their growth and performance, by discovering how to develop effective innovation strategies. Innovation is becoming a survival necessity in business and organizations. This work is concentrated in the Innovation research area, mainly, on how through AI, we could help automate innovation processes.

This paper carries out a review of works that use AI and intelligent systems to improve innovation capabilities in the MSMEs in the textile industry. The main questions of the systematic literature review are: How can the processes of innovation in MSMEs be improved? What intelligent systems have been developed to improve the innovation capabilities of MSMEs? Also, How AI has been used in the textile industry?

Vagnani & Volpe (2017) reviewed different meta-analysis approaches based on Structural Equation Modelling (SEM) to determine the associations between the attributes of innovations, the managers´ behavioral preferences, and the organizations' innovation adoption decisions. Zhong, et al., (2017) reviewed 165 papers considering concepts such as intelligent manufacturing, IoT-enabled manufacturing, and cloud manufacturing. They have used the Google Scholar and Scopus databases. Their paper discussed critical technologies, such as the IoT, cloud computing, Big Data Analytics (BDA), Cyber-Physical Systems (CPSs), and Information and Communications Technology (ICT) in the context of smart manufacturing. This work is different from previous works because it focuses on how intelligent systems have been used to improve organizational innovation capabilities, understanding the innovation process as a series of steps to achieve innovation in MSMEs.

The contributions of this article are: (1) the identification of how intelligent systems and AI have been used in the innovation processes of MSMEs; particularly in the textile industry, (2) a proposal for an innovation process based on the different innovation models that exist, designed to work autonomously in each of its phases, and (3) the analysis of current challenges in the application of AI techniques to improve the innovation processes in the MSMEs.

This document is structured as follows: in the next section, the theoretical framework about innovation models is addressed. In the third section, the systematic literature review protocol is presented. Section fourth presents the literature review´s main results. The conclusions are presented in section five, outlining the limitations of the study, the contributions to theory and practice, and the recommendations for future research.

CONCEPTUAL FRAMEWORK

The Innovation Models

Throughout history, different models have been developed to explain the process of innovation. The models that have been used the most are (Velasco et al., 2007; Nuchera, 2002; López, 2009): TRIZ, Design Thinking, and Lean Startup model. The following section is dedicated to explain the most used innovation models.

Triz Model

The knowledge management techniques provided by the Russian TRIZ methodology facilitate the development of systematic innovation at any level of an organization. Inventions range from a simple adaptation of something that already exists to an actual discovery that revolutionizes technology and science (level 1 to level 5 of inventiveness, according to the Altshuller classification) (Souchkov, 2010).

Figure 1 shows the TRIZ-based systematic innovation model, which response to the competitive business world´s modern needs. This model has been elaborated from two essential concepts: one of abstraction and the other of concretization.

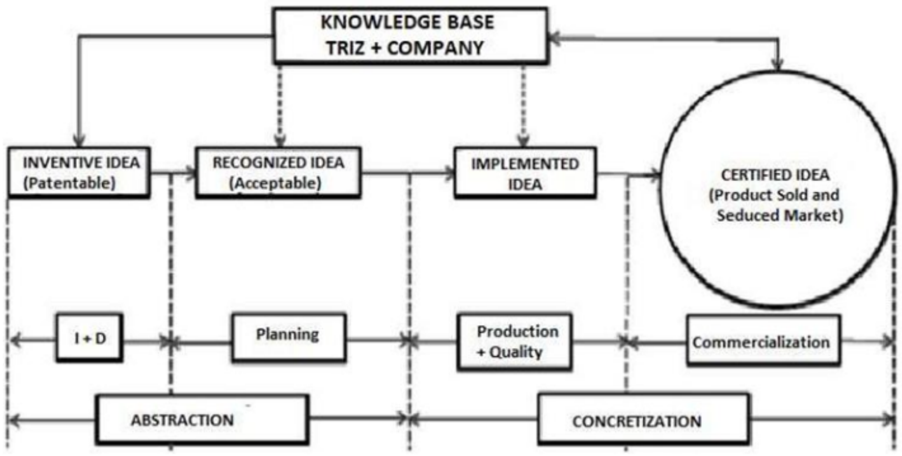

FIGURE 1
TRIZ-BASED SYSTEMATIC INNOVATION MODEL (SOURCE [9]).

In general, the conceptual model of Systematic Innovation with TRIZ covers two stages: abstraction and concretization. During the first stage, the idea is generated, and during the second stage, the idea is materialized in a product.

According to Souchkov (2010), the innovation process begins with a creative patentable idea and culminates when this idea is certified. The only way to certify an idea is the selling of a product based on this idea, which means that the primary indicator is the concretization. When concretization is achieved, the company´s knowledge base is evolving, and might become a source for future ideas.

In the first step, efforts are devoted to the research and development of an idea that must arise to solve any problem. The second step is the most complex, since the idea that has been generated must convince people responsible for approving or rejecting it. If the idea is accepted, it will be executed in the next stage, called concretization. The production and quality steps are crucial because a good idea could be poorly executed. Finally, the marketing step is essential, because a good product may not be sold due to wrong marketing. This last step is called "certified idea", which is the culmination of the initial abstract idea because the consumer certifies the idea with its loyalty.

In this model, the innovation process is systematic, and therefore, it is possible to foresee it, plan it, organize it, transmit it, delegate it, accelerate it and induce it as it is already being seen (Souchkov, 2010). Organizations that have used TRIZ, have not only survived the demands of the new market conditions, but made spectacular progress in their competitiveness with a significant impact on their profits, the well-being of the company and society in general (Souchkov, 2010).

Design Thinking

Design thinking is a methodology useful in solving ill-defined or unknown complex problems

(Friis, 2020; Interaction Design Foundation, 2020) This approach has five stages (see Figure 2) running through a process that refines an idea until reaching a solution that meets the team's objectives. These steps are described below.

Empathize: The Design Thinking process begins with a deep understanding of the needs of the users involved in the solution, and their environment to generate solutions consistent with their realities.

Define: During this stage, the information collected during the previous phase is screened to keep what adds value.

Ideate: The Ideation stage aims to generate endless options with activities that favor expansive thinking.

Prototype: This stage turns the ideas into reality; for this, prototypes are built to make the ideas palpable, highlighting elements that must be improved or refined before reaching the result.

Test: This phase carries out tests of the prototypes, with the users involved in it. This phase is crucial and helps identify significant improvements, failures to solve, and possible shortcomings.

Lean Startup Model

The Lean Startup Method refers to a philosophy or methodology used when addressing the launch of new businesses and products (Ries, 2008). Generally, it is applied in the technological sector, although its use extends to companies in other areas (Llamas & Fernández, 2018). Lean Startup activities are: Validate, Create a Minimum Viable Product (MVP), Launch, Measure, and Learn.

Hypothesis Validation: validate the assumptions from which the idea is derived, which must be confirmed by the market and the client.

Creation of an MVP: generate a product with minimal characteristics that add value to the client.

Metrics: establish quantifiable elements to measure the expected performance with the new product.

Launch and Learn: the product is launched, and the learning obtained through iterations due to the utilization of the product will allow continuing improvements, and detecting necessary changes.

Proposed Model for the Innovation Process

An innovation process is a structured strategy for ensuring that the innovation team idealizes an innovation and progresses on it until it is successfully implemented. This section explains a proposal of an Innovation process based on the different innovation models that exist (see Figure 2). In the proposed approach to the innovation process, four macro-processes are considered: Problem Analysis, Ideation, Experimentation, and Commercialization.

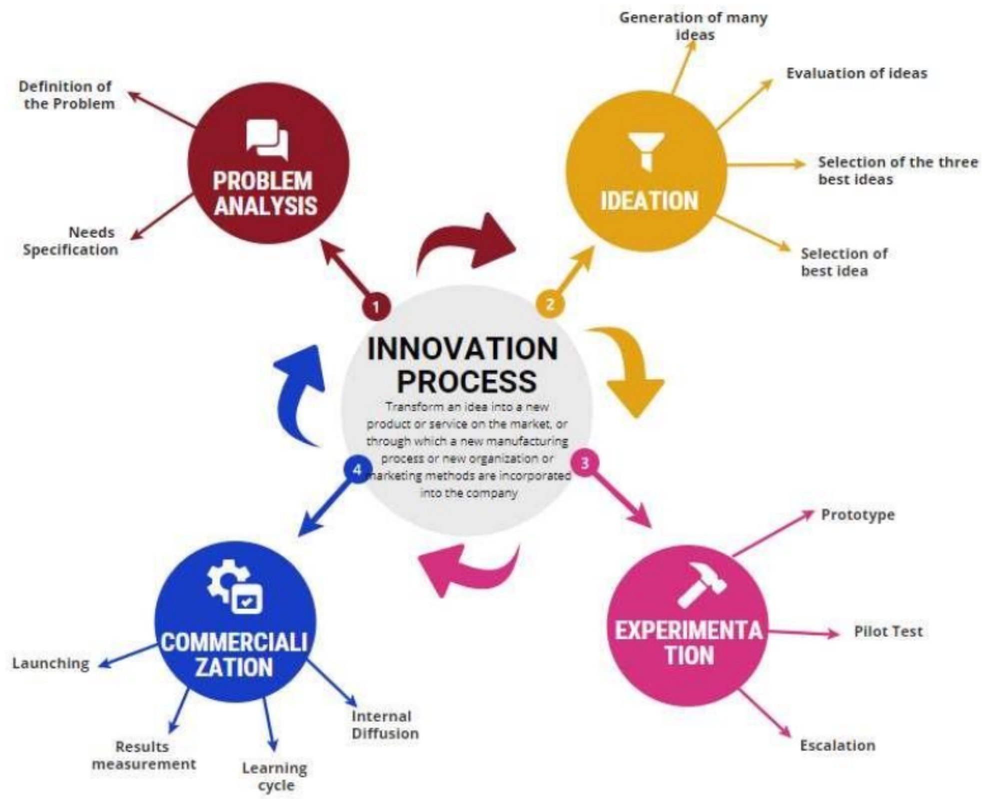

FIGURE 2
PROPOSED INNOVATION PROCESS MODEL

1. **Problem Analysis:** The problem must be identified and defined.
    1.1 Definition of the problem: in this step, it must be indicated and defined what the problem is.
    1.2 Needs Specification: it defines a list of requirements necessary to solve
2. **Ideation:** Different ways to solve the problem must be explored.
    2.1 Generation of ideas: In this step, several ideas are generated. The amount here matters. The more, the better. It can use the technique of Brainstorming.
    2.2 Ideas Evaluation: it is the process of comparing and contrasting ideas related to the new product, to select the most promising.
    2.3 Selection of the three best ideas.
    2.4 Choice of the best idea: The idea that best solves the problem.
3. **Experimentation:** A palpable version is generated, although not necessarily exact to the initially proposed product.
    3.1 Prototype: It is the development of an initial product, which allows deciding if it is feasible.
    3.2 Test: The main objective is to validate the creative process
    3.3 Escalation: Means to transform a concept (prototype) in a commercial product
4. **Commercialization:** This is the process of launching new goods or services to the market.
    4.1 Launching: it is oriented to publicize the innovation product and its results.
    4.2 Results measurement: metrics to measure the results should be defined.
    4.3 Learning cycle: The market will give feedback to know if the product must be changed, optimized, or persevere.
    4.4 Internal Diffusion: it is the communication to internal clients and employees. The objective is the utilization of innovation as a positive reinforcement to motivate individuals within the organization.

Table 1 shows the relationship between the Innovation models explored and our proposed Innovation process. In general, each of the models contributed effectively to the phases and sub-processes of the proposed approach. The Lean Startup models and the Triz Model include all the macro-processes defined for the proposed model.

Table 1
MAPPING OF INNOVATION MODELS IN THE PROPOSED INNOVATION

Innovation Process/Model	TRIZ model	Design Thinking	Lean Startup
Problem Analysis	✓	✓	✓
Ideation	✓	✓	✓
Experimentation	✓	✓	✓
Commercialization	✓		✓

The differences between the proposed model and the other models are as follows:

Process

The differences between the proposed model and the other models are as follows:

- **The Starting Point is Different:** In the Lean Startup model, the process starts with an idea, which must be validated and can vary significantly before reaching the market. In Design Thinking, the starting point is a challenge (problem, need), not an idea. In our model, the starting point is the process of analyzing the problem, which begins with the definition of the problem and the specification of the needs of the identified problem.
- **The Terminology is Different to Express Similar Concepts:** For example, the Lean Startup model talks about MVP, the Design Thinking model talks about the prototype, while our model talks about the experimentation process with the prototype, testing pilot and escalation phases.
- **The Number of Phases is Different for each Model:** Particularly, the Lean Startup model has four phases, our innovation model also has four phases, but some of the phases in the Lean Startup model are included in our phases.

The main contributions of this approach are that each of the phases of the model is self-contained, autonomous, and can be self-managed. The characteristic of the phases described will allow the adequate definition of autonomous cycles of data analysis tasks (Sánchez et al., 2016) in future works. This model facilitates automatically to reach the goals of each phase by using the data from the innovation process and AI techniques for their analysis, as will be supported by the following literature review.

DESCRIPTION OF THE PROTOCOL FOR THE SYSTEMATIC LITERATURE REVIEW

The first step to initiate the systematic review of the literature was to define the objective of the review, which was to establish the state of art in relation to intelligent systems that improve the innovation capacities in MSMEs. In the following steps, the procedure for the systematic literature review is described.

Research Question

The research questions defined around the objective of this systematic literature review are defined in Table 2. These questions are about the innovation processes in the MSMEs, the utilization of intelligent systems and AI for the innovation in the MSMEs, with a focus on the textile industry.

Table 2 LIST OF SEARCH QUESTIONS	
Question	**Description**
Q1	How can the processes of innovation in MSMEs be improved?
Q2	What intelligent systems have been developed to improve the innovation capabilities of MSMEs?
Q3	How AI has been used in the textile industry?

Inclusion and Exclusion Criteria

The inclusion criteria considered to apply filters in the article search are listed in Table 3.

Table 3 LIST OF INCLUSION CRITERIA	
Criteria	**Description**
IC-1	Include publications with titles related to intelligent systems used in MSMEs, emphasizing their application in innovation processes
IC-2	Include posts that contain keywords that matching those defined in the search string.
IC-3	Include studies in two languages: English and Spanish
IC-4	Include publications from the last five years

The exclusion criteria that were defined for the systematic review of the literature are shown in Table 4.

Table 4 LIST OF EXCLUSION CRITERIA	
Criteria	**Description**
EC-1	Exclude publications that do not match the previous inclusion criteria.
EC-2	Exclude all duplicate posts.
EC-3	Non-digital publications (publications not available for review)

The digital databases taken into account in the review to perform the search were ACM, WoS, Scopus, Google Scholar, IEEE Xplore, Science Direct, Wiley, Springer.

Search Chains

For the definition of the search chains (see Table 5), the keywords are:

- Linked to AI: "data analytics", "artificial intelligence", "data analysis",
- Linked to the organizations "MSMEs", "textile industry", "Smart MSMEs",
- Linked to innovative processes: "Innovation Ecosystem", "innovation strategies", "innovation processes" Quality criteria.

Table 5 SEARCH STRINGS PER RESEARCH QUESTION		
Question	**Description**	**Search string**
Q1	How can be improved the processes of innovation in MSMEs?	("smart innovation capabilities" OR "innovation processes" OR "Ecosystem innovation") AND ("micro small and medium business" OR SME OR MSME)
Q2	What intelligent systems/models have been developed that improve the innovation capabilities of MSMEs?	("intelligent system" OR "intelligent model" OR "Artificial Intelligence" OR "data analysis" or "data analytic") AND ("innovation capabilities" OR "innovation processes" OR "innovation processes in MSMEs" "Ecosystem innovation" OR "innovation processes in SMEs")
Q3	How has AI been used in the textile industry?	("intelligent system" OR "intelligent model" OR "Artificial Intelligence" OR "data analysis" or "data analytic") AND "textile industry"

Search Strategy

For the development of the protocol, preliminary searches were conducted aimed to identify existing systematic literature reviews and assess the volume of potentially relevant studies. To obtain the final list of publications to analyze, several combinations of search keywords for each research question were used. According to the searches carried out, new keywords were added. The process of selection of articles was carried out using the search chains defined for each research question, and applying the exclusion and inclusion criteria, according to the following steps:

Step 1: Search Without Filters: it defines the total number of articles for each search question, without taking into account any type of inclusion or exclusion criteria

Step 2: Publications from the Last Five Years: the temporality inclusion criterion was applied, and articles with a publication date that was within the last five years were taken into account.

Step 3: Type of Paper: this filter excludes papers that are not in Journals or conferences.

Step 4: Keywords and Abstract: this filter is applied because the search chains must be contained in the keywords and the article´s summary.

Step 5: Title: this filter is applied because the search chains must be contained in the article´s title.

Step 6: Final Selection: this step selects the articles from the previous step, which match the search questions.

Figure 3 shows the flow chart of the selection process, and finally, 44 articles were suitable for proceeding with the analysis.

FIGURE 3
FLOW DIAGRAM OF THE ARTICLE SELECTION PROCESS

Preliminary Results

Figure 4 shows the distribution of the production of selected scientific papers per country. Overall, China and Italy are the countries with the most papers with 9%. Also, Germany and

Thailand seem to have a solid tradition to study topics related to innovation capability in small businesses with 7%. According to the World Innovation Index 2019, the innovation in China is evident in numerous spheres; the country maintains first place in patents by origin, industrial designs and trademarks by origin, net exports of high technology, and exports of creative products. Thus, China is firmly installed in the group of leading innovative nations. In Italy´s case, the investment on research and development reaches only 1.38 percent of GDP (Gross Domestic Product), a number that despite its growth in recent years remains low. However, Italy occupies the eighth place in the world in scientific publications (second in our domain), an issue that has been discussed by Mauro Battocchi, ambassador of Italy in Chile (Siebert, 2020): "The academic world is working well to create science and research, and we see it in publications; Companies are doing a good job, and after all Italy is well-positioned to face new challenges, but the biggest issue is technology transfer. If we can combine successful scientific production with the strength of the Italian company, Italy will be even stronger". This situation is similar in Germany, the leader in industry 4.0, and in Thailand, where the nation has experienced massive economic growth for the past years, based on its innovation performance.

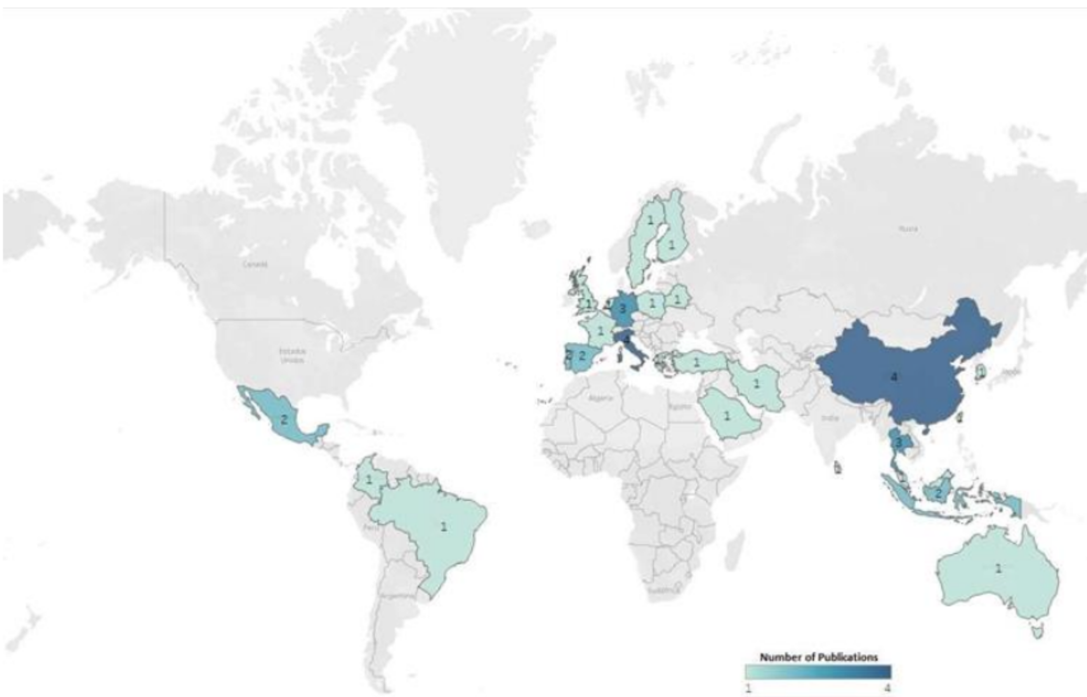

FIGURE 4
SCIENTIFIC PRODUCTION PER COUNTRY

In figure 5, the word cloud of all the articles selected to perform the review is presented. Those words with the highest repetition number correspond to: innovation process (19), innovation (15), innovation capability (18), micro small and medium business (7), IA Textile (11). These words were considered in our search string within the protocol.

FIGURE 5
WORD CLOUD

The most frequent keywords have been used to generate a Sankey diagram (see Figure 6), which helps visualize the technologies, methods or processes considered in the reviewed papers.

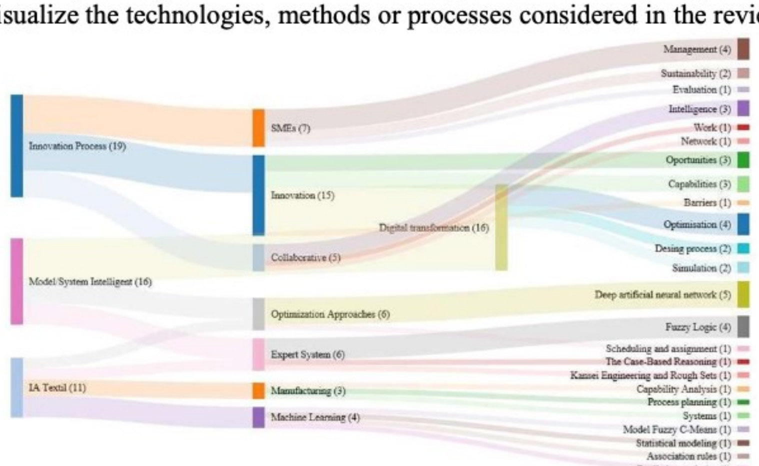

FIGURE 6
SANKEY DIAGRAM FOR KEYWORDS

According to Figure 6, medium and small companies try to evaluate and improve their processes by implementing strategies in the context of management and sustainability. Thanks to the development of innovation in those organizations, they have identified the barriers, but they have also defined opportunities and capabilities through collaborative work, in a network based on digital transformation. Finally, the articles showed how AI has been applied in the textile industry (e.g., optimization, simulation) through different techniques (e.g., deep learning and fuzzy logic).

ANALYSIS OF THE PAPERS

Improvement of the Processes of Innovation in MSMEs

Young & Seong (2016) investigated ways to maximize firm performance according to design Innovation management and market dynamics. They defined a multi-agent simulation approach to understand the MSMEs´ performance, considering each MSME as an agent and observing their innovation activities. They concluded that the level of firm diversity influences the performance of MSMEs' innovative activities. They designed an Agent-Based Innovation Simulator to experiment based on the following assumptions: (1) every MSME attempts to execute innovative activities to accomplish its work efficiently; (2) each MSME prefers to work with those partners (MSMEs) in a similar category; (3) a single MSME tends to be divided into a number of subgroups; and (4) diversity in how MSMEs influences the firm's decision to collaborate with others in order to complete their work effectively.

Niewöhner et al. [13] identified challenges to which companies must respond as part of their innovation management to obtain a holistic picture of all innovation management activities; they have defined a total of six fields of action that must be taken into account in connection with innovations (see Figure 7).

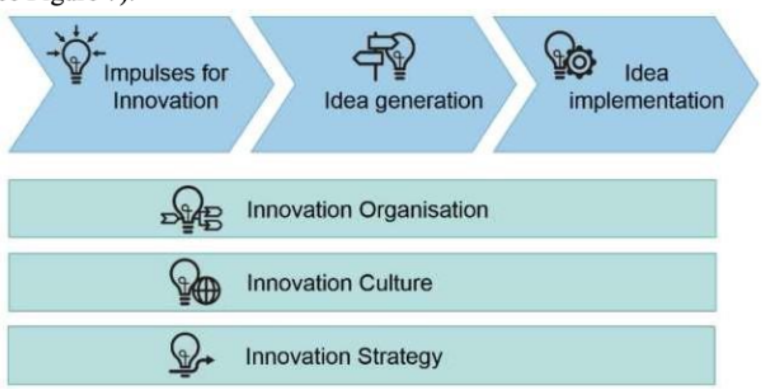

FIGURE 7
THE SIX DESIGN STEPS FOR INNOVATION MANAGEMENT (SOURCES [10]).

- **Impulses for Innovations:** it comprises all activities that ensure the fundamental knowledge generated, from which ideas can be developed
- **Idea Generation:** it is about generating innovative ideas. Based on the knowledge of the domain, concrete ideas are generated
- **Idea Implementation:** it involves the successive concretization of an innovative idea
- **Innovation Organization:** the central question in the field of innovation is what must be the structural conditions to carry out innovation projects successfully.
- **Innovation Culture:** a successful innovation process is not only based on knowledge and experience, but requires an appropriate environment in which innovative ideas can grow.
- **Innovation Strategy:** the innovation strategy is necessary to ensure that innovation activities are not without purpose.

Oleśków, et al., (2017) analyzed the innovativeness levels in companies and industries of various sizes of the entire world, and concluded that the innovativeness of companies is at low to moderate level. Additionally, they proposed an open space platform, where organizations can evaluate their innovation competencies, which allow cross-sectorial and cross-border cooperation. Finally, they proposed roadmaps about projects and organizations, to help industrial development and generate new industrial value chains.

Bagheri, et al., (2019) examined 116 SMEs from eight sectors that showcase high innovative activity within the UK, and offered new insights into internationalization´s role in their international performance. The fundamental question in this study was why and how technological innovation determines the international firm's performance. Technological innovation plays a crucial role in achieving this competitiveness and increasing firm performance. The model predicts that internationalization can improve firm performance, whereas technological innovation has a mediating role.

Rifat et al. (2015) described how the smaller firms´ coexistence with larger companies, in recent years, helped remove innovation limitations and achieved the required performance in global competition. Moreover, vertical and horizontal cooperation with customers, providers, and other participants, play a significant role in small and medium organizations´ innovation processes. On the other hand, Rahab (2015) analyzed the relationship between innovativeness and market orientation. They concluded that there is a positive influence of the market orientation on innovativeness, such that the firms' innovativeness level increase when they pay attention to the market orientation.

The research of Rifat, et al., (2016) provided new evidence on the effects of KMC (Knowledge Management Capabilities) on innovation performance in the emerging market of Turkey, where fast information and knowledge integration, along with the application, are crucial for firms to adapt to a dynamic environment. In this research, the authors developed a questionnaire for the largest 1000 firms in Turkey from different sectors. The questionnaire consisted of 41 questions: 12 questions to measure the performance effects of KMC, five questions for ED (Environment Dynamism), four questions for innovation performance, 11 questions for LC (Learning Capability), five questions to control the effects of industry structure factors, one question for the stage of industry's life cycle, two questions for demographic aspects (age and size), and the last question to categorize the primary business activity of the firms. Findings highlight the importance of KMC and LC, and the critical role of knowledge management in innovation performance.

Xing, et al., (2019) explored the relationship between GOI (Green Organizational Identity) and CEP (Corporate Environmental Performance) in a moderated meditation model, which includes SER (Sustainability Exploration Innovation), SEI (Sustainability Exploitation Innovation), and GER (Government Environmental Regulation). Using SEM and a bootstrap method based on data sets from 380 Chinese companies, they conclude the moderating role of GER, and the indirect effect of GOI on CEP via SER and SEI, revealing the driving effect of GER on innovation and performance.

Xin & Sun (2016) carry out an empirical study to explore if innovation offsets the corporation loss caused by environmental regulation and to improve corporation competitiveness. They analyze Chinese pollution-intensive corporations during 2007–2012, and the results showed that the implementation of flexible control policies had achieved initial success and environmental regulations had significant positive effects on the corporation's innovation.

Maletič, et al., (2015) addressed the gap of sustainability-oriented innovation practices, which result in significantly improved products, services, processes and management systems. The study analyzed the effects on the performance dimensions (i.e., economic performance, quality performance, innovation performance, environmental performance and social performance), considering manufacturing and service industries across five countries: Germany, Poland, Serbia, Slovenia and Spain. The regression analysis results demonstrated that sustainability- oriented innovation practices are positively related to overall organizational performance. From a practical perspective, the study explained how organizations can embed sustainability aspects in their innovation processes with the aim to improve their overall performance.

In their study, Conesa, et al., (2016) assessed the relationship of economic sustainability with organizational innovation and firm performance in a single integrated model, by using SEM in a dataset of 552 Spanish firms. The authors proposed a model of the relationship between the three dimensions of sustainability: CSR activities (social and environmental sustainability), innovation performance and firm performance (economic sustainability) in Spanish SMEs. The results supported a partial mediation effect of innovation performance on the relationship between CSR and firm performance, explaining how CSR is an important driver mechanism for companies to be more innovative, efficient and effective.

Pittawat & Alisara (2017) aimed to help improve the innovation of Thai SMEs in food manufacturing. They analyze the triple helix model of innovation that describes the interactions between government, academia, and industry, to promote economic and social development, by considering 87 responses of a questionnaire regarding SME collaboration. The majority of the samples were fruit and vegetable processing manufacturers, followed by meat and fish/aquatic animal manufacturers. Multiple regressions was applied to quantify the relationship between the identified factors, and findings showed that SME collaboration with the industrial sector was at the highest level when compared with the university and government sectors, respectively. Also, the results indicate that SME collaboration with the industry sector had a positive effect in the innovation of Thai food manufacturing SMEs, and the government sector also has a positive influence.

Apanasovich, et al., (2016) analyzed the most effective innovation modes in the context of post-Soviet Transition Economies (PSTE), specifically in Belarus. In the study, the study considered two dependent variables: product innovation and organizational innovation. By applying two regression models, one for the product innovation output (ProdIO), and the other for organizational innovation output (OrgIO), results demonstrated that there is a positive correlation between innovation modes and product innovation.

The purpose of the paper of Castela, et al., (2018) was to propose an innovative management methodology to evaluate innovation capability in SMEs, through the combined use of cognitive mapping and the Analytic Hierarchy Process (AHP). They create a non-parametric method for the evaluation of SME innovation capabilities. The proposed approach is useful to evaluate SME innovation capability because the collective cognitive map helps to structure well the problem.

De Martino & Magnotti (2018) analyzed the innovation capacity of agrifood MSMEs in the Campania region, through a theoretical framework that addresses the influence of some external and internal resources on the development of innovation activities. These variables are about the local socio-economic external context of firms, the internal resources (investment in R&D, number, competencies and revenues of employees, etc.), as well as the collaborative relationships with externals partner (research organizations, suppliers, etc.). The article remarked the significant role of public funding in boosting the innovation capacity of agrifood MSMEs; this is a critical element for developing collaborative innovation, which at the regional level has to be performed by public decision-makers, both as a promoter or initiator of innovation networks. The implementation and the creation of innovation networks, such as the case of the technological districts, or the development of research projects can generate and favor the circulation of knowledge; thus, increasing the innovation capacity of agrifood MSMEs.

Kritsadee, et al., (2017) tested a model of factors affecting the innovativeness of SMEs using SEM. Data about innovativeness were collected using questionnaires that were mailed to 283 entrepreneurs. The proposed model determined that learning orientation and proactiveness had direct effects on innovativeness. The results are a guideline to build innovativeness in enterprises. The analysis addressed product, process, as well as organizational and marketing innovation, and the contribution of them to the organization's results, e.g., market share, environmental sustainability, profit, etc.

Based on a survey from Korean SMEs in the manufacturing sector, Kim, et al., (2018) established that innovation capability is a fundamental aspect of firm performance, and the organizational rigidity and insufficient resources are barriers. Finally, they concluded that top management leadership and external networking impact innovation capability. Guo, et al., (2019) divided enterprises´ external knowledge acquisition into external technical knowledge acquisition and external market knowledge acquisition, and examine this theoretical framework with the environmental innovation as a mediator. They used a sample of 416 Chinese SMEs, and the results revealed that high levels of technical and market knowledge acquisition are related to SMEs' performance the positive impact of technical knowledge acquisition is mainly embodied through economic performance; while market knowledge acquisition´s positive impact is mainly embodied through environmental performance. Meanwhile, environmental performance plays a mediating role between environmental innovation and economic performance; and environmental innovation plays a mediating role between external knowledge acquisition and enterprise performance.

Guzman, et al., (2018) investigated the effects that innovation capabilities have on the business performance of 308 SMEs within the context of Aguascalientes, Mexico. The results showed a positive effect on Mexican SMEs´ performance of the innovation in products, processes, marketing and management have. Thus, this paper studied the relationship between innovation capabilities and business performance, and defined some of the innovative aspects to consider in order formulating innovation strategies.

Table 6 describes the domains in the innovation process context where the previous works were applied (e.g., for the management of innovation processes or to analyze the innovation capability).

Table 6
DOMAIN IN THE INNOVATION PROCESSES WHERE THE PREVIOUS WORKS HAVE BEEN APPLIED

Domain	Meaning
Innovation management	In the context of idea generation, Idea implementation, organizational Innovation, and Innovation culture
Innovation in the SMEs	It is related to product, process, marketing, or organization
Sustainability Impact in the innovation processes	Improvement in economic performance, quality performance, innovation performance, environmental performance, and social performance
Method to evaluate the Innovation capacity	Multivariate statistical methods; forecasting methods, Uncertain linguistic variables, fuzzy logic

Intelligent Systems to Improve the Innovation Capabilities of MSMEs

Giulia (2016) proposed an innovation management system to analyze the data and classify the possible changes in the production process to allow new products. They propose changes in the components or in the production operations. Gasparetto, et al., (2018) proposed a framework in which an Intelligent Work piece Carrier (IWC) is introduced in a traditional production line for a flexible organizational model. This work integrates the knowledge of production steps in the IWC to allow decision making about the process execution. They develop a first prototype to verify the effectiveness of the proposed framework and the results showed that the IWC is a necessary component for flexible production systems.

Cluster manufacturing has been used in the world to improve competitiveness via sharing resources and profits. Suthep & Chiwapon (2019) presented a model for performance evaluation of the bus body cluster manufacturing in Thailand. Fuzzy logic is used to determine the efficiency of each factory according to five dimensions: Reliability, Flexibility, Source, Capacity, and Quality, providing a cost optimization for each situation.

Singh, et al., (2016) proposed a fuzzy rule-based expert system for manufacturing performance assessment in SMEs. The sixteen used metrics were categorized as four economic, five environmental and three social measures. They propose to gather the inputs in terms of linguistic variables in the decision-making process of manufacturing SMEs. The fuzzy rule-based expert system elicits all the aspects of the organization based on the triple-bottom-line framework. On the other hand, Bilal, et al., (2019) propose a system to evaluate the product development process. The system helps decision-making in large enterprises manufacturing, or a group of SMEs, by using explicit knowledge of formal decision events. It stores and reuses the events and experiences related to different activities of the industrial production process, (e.g., product design, manufacturing, and product inspection).

Applications using Artificial Neural Networks (ANN) have been also revised. For example, Silva, et al., (2019) compared the performance of several ML techniques like ANN, Support Vector Machines and Support Vector Regression (SVM-SVR) machines. They demonstrated the importance of considering external factors, such as macroeconomic and microeconomic indicators, like the prices of related products that affect the level of sales in an organization. Zygiaris (2017) conceptualized an information extraction mechanism for semi-structured data, based on web semantic ontologies, in order to introduce an initial stage of the innovation intelligence chain. They use an ontological framework related to innovation chain intelligence, in order to extract the information, and represent and update the knowledge.

Also, innovation capability is a key factor s for surviving in a hyper-competitive situation. The study of Allawi (2018) described the influence of knowledge sharing in innovation capability. The study used a hybrid (qualitative and quantitative) approach for an effect-causal analysis. The factual analysis was conducted based on a survey of 145 employees in a telecommunication company in Indonesia. They concluded that knowledge sharing has a positive influence on creativity, as well as creativity has a positive influence on innovation capability.

The digital transformation is forcing organizations to go towards data-driven business models. Ossi, et al., (2019) presented a conceptual framework to explain the innovation capabilities based on big data and business models. Using the design science research method approach, the authors built the framework based on the existing literature and applied it in three firms.

Adamides & Karacapilidis (2020) proposed an ICT support framework focused on the micro- foundations of dynamic capabilities: “opening” mental models, and development of social capital through collaboration and knowledge management; while the aim of Álvarez-Castañón (2019) was to analyze the dimensions of the Open Innovation (OI) and the technological and innovation capabilities of SMEs in a traditional manufacturing industry. A regression analysis was conducted in a sample of 33 SMEs in Guanajuato, Mexico.

Table 7 describes how IA techniques have been used in the innovation processes in previous works.

Table 7
SOME INTELLIGENT SYSTEMS IN THE CONTEXT OF INNOVATION PROCESSES

IA method	Objective
Fuzzy logic	To optimize the efficiency of each company considering five dimensions, which are capacity, quality, reliability, source, and flexibility Expert system based on the characteristics of SMEs
Artificial Neural Networks	For product sales forecasting with high rotation
Semantic Web ontologies	For information extraction and knowledge representation from the production chain

The AI in the Textile Industry

The textile industry has developed expert systems applications to increase production, improve quality and reduce costs (Weldeslasie & Ahmed, 2019). The analysis of textile designs or structures, includes the use of mathematical models like the Finite Element Method (FEM), which has facilitated the prediction of the behavior of that textile structure under mechanical loads (Arjmandi et al., 2017). Finally, the Case-Based Reasoning (CBR) method proposed by Bullon, et al., (2017) complemented those previous´ methods results.

The paper of Santiago-Santiago, et al., (2017) presented an interactive design support system for fashion designers when designing workwear or corporate wear clothes. This system is fed by a semantic database that describes the relations between function and clothes in a specific context of use. This application contains relevant information for clothes designers and producers, alerting them with user’s clothes preferences for a particular task.

On the other hand, Shih-Wen proposed a predicting model to identify color trends (Shih-Wen et al., 2016). The method, based on fuzzy c-means, was used to separate the collected color data, and the minimum mean-square error was then used to determine the color clusters. In order to verify the quality of the system, four datasets of Pantone from spring 2014 to fall 2015 were taken as samples, and the color for spring 2016 was predicted and compared with that of Pantone spring 2016. The work of Vinod & Poonkuzhali (2018) focused on sales forecasting using trend analysis of garment industries based on the e-commerce data collected from the Coimbatore region, a region in India of cotton production with a high density of textile industries. Time series analysis is performed, and the forecasting about sales trends is derived from the data.

Zimmerling, et al., (2019) investigated the utilization of Machine Learning (ML) techniques in the textile context, using variable geometries to obtain physically accurate predictions of images from the textile industry, recognizing distinct geometric features in a textile piece using ML- algorithms. They analyzed the algorithm’s learning capacity vis-a-vis geometry variations. This work used fabric forming for approach validation, which is applied to other material classes (for example, stitched non-crimp fabrics).

The purpose of Klement, et al., (2017) was to develop a tool to solve scheduling problems in a textile company, in order to assign the resources through time. To do so, they use a hybridization between a metaheuristic and a list algorithm. The metaheuristic can be used in any problem, and the list algorithm is very specific to the studied problem.

Optimizing the parameters of the manufacturing process requires an intensive search for resources in a high-dimensional parameter space. Pfrommer, et al., (2018) proposed a substitute-based optimization as a simplified model to guide the search for optimized parameter combinations, where the substitute model was iteratively improved with new observations to find optimized parameter settings. They applied substitute-based optimization in a composite textile draping process. In addition, a deep neural artificial model was trained to predict the cutting angle of more than 24,000 textile elements. Finally, Packianather, et al., (2017) proposed data mining techniques applied to a manufacturing SME, like the ant colony model that has been largely used to solve combinatorial optimization problems (Aguilar, 2001) or the artificial immune systems to solve detection problems (Araujo et al., 2003).

Murphy, et al., (2019) investigated the use of ML technologies for order flow time estimation in dynamic job shops, using a discrete event simulation framework for modeling manufacturing operations. Data was generated via simulation, a series of experiments were conducted, and the proposed approach´s performance was compared with conventional methods. In this paper, five ML models were tested: ANNs, gradient boosting, and regression trees with the ensemble and boosting concepts. Complex dynamic shop behaviors, such as rework, machine downtime, machine setup time, and machine repair time were considered. A series of simulation runs are then conducted with varying dispatching rules, such as FCFS (First Come First Served), SPT (Shortest Processing time first), EDD (Earliest Due Date first), to collect data for training, validating and testing of the ML models.

Aiming at the current requirements of the manufacturing execution system of SMEs, and with the development of enterprise in cloud manufacturing, Zhou, et al. (2019) proposed a manufacturing cloud platform application research for SMEs based on micro-service architecture. In this research, they used micro-service architecture to implement a platform- based business system, which solves very well the scalability problems of business applications and catches user requirements.

Table 8 shows the different AI methods used in the context of the textile industry. In general, different ML techniques have been used, but other techniques also. Besides, there are different types of applications of those methods (e.g., to support the design process, to improve the manufacturing process).

Table 8
AI METHODS USED IN THE TEXTILE INDUSTRY

AI technique	Manufacturing processes in the textile industry	Ref
Expert System	An expert system for production.	Bullon et al., 2007
	An expert system for clothing design processes	Santiago-Santiago et al., 2015
Machine Learning	An intelligent system for fashion color prediction	Shih-Wen et al., 2016

	A forecasting market trends approach to determine	
	the sales of garments	Vinod & Poonkuzhali, 2019
	Textile formability assessment	Packianather, et al., 2017
	Estimation of the order flowtime	Zimmerling, et al., 2019
Optimization	Solving scheduling problems	Packianather, et al., 2017
Approaches	Parameters optimization of a manufacturing process	Zimmerling, et al., 2019

DISCUSSION

Improvement of the Innovation Processes in MSMEs

The Triple Helix concept has been discussed in the development of dynamics of knowledge based on innovation systems (Pittawat & Alisara, 2017). Recent studies have explored the relationship between SMEs and this concept to drive innovation performance. One challenge is to explore the Triple Helix, and even the Quadruple Helix concept (Carayannis & Campbell, 2009) (which includes the civil society as a participant) to the fashion sector, taking into account the benefits that MSMEs would have in terms of collaboration, resource access knowledge transfer and innovation skills, which could be applied in the innovation processes of MSMEs. Also, the creation of industrial value chains contributes to regional smart specialization strategies by capitalizing on complementary competencies, which positively impacts the innovation performance of MSMEs (Oleśków et al., 2017).

MSMEs should be open to learn new things via different activities, such as market surveys, observation and listening of customer comments and innovation related training. These activities allow increasing the knowledge to develop new creative ideas, meeting the customer demands. One challenge is to automate the process of generating creative ideas with the sources of knowledge already available in the market and the organization.

Future research should help the design, development and implementation of frameworks and systems to help organizations, especially MSMEs, in their innovation processes, as MSMEs engaged in more innovation processes present better overall international performance and competitive advantages (Bagheri et al., 2019). As seen in the literature review, intelligent systems and AI might help the innovation process; therefore a challenge for future research and development is to work on computational frameworks to support the innovation processes of the MSMEs.

On the other hand, organizations might consider sustainability aspects in their innovation processes, not only for corporate responsibility, but for the positive impact on their overall performance. One research opportunity then is to examine the factors that drive the sustainability-oriented innovation practices in organizations, especially in MSMEs. Thus, it is necessary to better explore the variables that moderate the relationship between sustainability- oriented innovation practices and organizational performance, understanding the importance of differential innovation attributes and organizational characteristics.

Intelligent Systems to Improve the Innovation Capabilities of MSMEs

Knowledge creation and management have proven to be very important for the innovation process, and they require the use of analytical tools and benefit from the use of intelligent systems and (AI Zygiaris, 2017). It is a managerial challenge to implement systems that allow refining the innovation process with the use of data analysis tools, to offer the possibility of improving the innovation capabilities of MSMEs.

MSMEs are a constant source of data valuable for innovation, concerning products, processes, management and marketing, but in many cases, they do not have systems that allow them to manage this information efficiently. One managerial challenge is to build data repositories with an adequate data analysis framework that exploits these sources of data to help in the decision- making process.

The use of IA techniques allows automating activities like decision-making, problem-solving, and learning. A challenge is to use them in the context of MSMEs, to transform these organizations towards data-driven business models. Finally, it would be valuable for theoretical and practical considerations, to evaluate the innovation capability of SMEs using automatic tools based on IA techniques like fuzzy cognitive maps (Aguilar, 2013) or multiagent systems (Aguilar et al., 2005).

The AI in the Textile Industry

As indicated by the literature review, the textile industry might use expert systems applications for product innovation, business expansion and user requirement, and to increase production, improve quality and reduce costs. Also, personalization is a challenge for the textile industry, and IA can be used to determine new products or characteristics, based on the profile of the clients (Reza et al., 2018).

The demand for personalized products makes necessary flexible production solutions that can deal with the problems of changing environments (Torn & Vaneker, 2019). Future research should help the identification and definition of smart environments for decision making that increase productivity without the need for the human component, in order to do the production processes more autonomous, personalized and efficient. Another challenge is the creation of an intelligent agent for the textile industry, which can learn innovative strategies and share them with other members of the organization.

The textile industry generates huge amounts of data as a result of thousands of transactions and processes that are carried out daily in all points of the value chain (Constantinescu et al., 2015). The challenge is how to intelligently manage all that volume of information that is being generated, using a smart data warehouse based on the Big Data and Data Analysis paradigms, in order to use this data to make better-informed decisions (e.g., in the operational and marketing areas) (Sivalogathasan, 2016; Pedron et al., 2018).

The possibilities for the utilization of AI are varied and wide, but if AI technology is included in the industry, then experts in this field will be needed, and collaboration with the scientific sector will be crucial. Also, the quadruple helix model will be necessary, especially, the collaboration of the academy with experts from companies in the same field. Also, refinement collaboration between companies will be crucial, as the more information we have the more effective and secure the developed models will be. AI will move the technological innovation systems.

Crossed Challenges

According to the systematic literature review performed, there are no studies related to features engineering that allow the analysis of the variables required in the innovation processes of MSMEs. A feature engineering work will allow identifying the most relevant variables that describe and influence the innovation processes. Also, future research should focus on exploring and defining the most appropriate AI technologies so that the MSMEs improve their innovation processes.

In the globalization context, it is important to consider the influence that the international markets have on MSMEs to improve their innovation strategies. Future research should address the incorporation of data analysis techniques in the innovation process, to gain knowledge about international markets and their needs.

The results of the previous studies demonstrate the importance of knowledge management within organizations and its impact on the innovation process and innovation performance, consequently, it is necessary to create tools for optimal knowledge management, particularly the use of analytical tools to support decision-making during the innovation process is a big challenge, in theoretical and practical terms.

ACKNOWLEDGEMENT

The authors thank Ministry of Science, Technology and Innovation and Universidad EAFIT for Bicentennial Scholarships that helped support this work.

REFERENCES

Ferrer, L.E., González, K.P., & Mendoza, L.M. (2015). Innovation as a key factor to improve the competitiveness of SMEs in the Department of the Atlantic, Colombia. *Free Opinion, 16,* 21-36.

Sánchez, O.M., Cervantes, A.V., & Peralta, M.P. (2016). Management of innovation in small and medium-sized companies in Barranquilla-Colombia. *Journal of Social Sciences, XXII*(2), 79-84.

Flórez, N.L. (2019). *Innovation as management strategies for SMEs.* Master Thesis, Nueva Granada Military University, Colombia.

Vagnani G., & Volpe L. (2017). Innovation attributes and managers' decisions about the adoption of innovations in organizations: A meta-analytical review. *International Journal of Innovation Studies, 1*(2), 107-133.

Zhong, R.Y., Xu, X., Klotz, E., & Newman, S.T. (2017). Intelligent manufacturing in the context of industry 4.0. *Review Engineering, 3*(5), 616–630.

Velasco, E., Zamanillo, I., & GurutzeIntxaurburu C. (2007). Evolution of the models on the innovation process: From the linear model to the innovation systems. *In XX Annual Congress of AEDEM, 2.*

Nuchera, A., León, G., & Pavón J. (2002). *The management of innovation and technology in organizations.* Editorial Pyramid. Madrid.

López, O., Blanco, M., & Guerra, S. (2009). Evolution of innovation management models. *Negative Innings, 5*(2), 251-264.

Souchkov, V. (2010). TRIZ and systematic business model innovation. *In Global ETRIA Conference TRIZ Future 2010.*

Friis, R., & Siang, Y. (2020). 5 stages in the design thinking process.

Llamas F.J., Fernández, J.C. (2018). The Lean Startup methodology: development and application for entrepreneurship. *Revista EAN, 84,* 79-95.

Young, W., & Seong, W. (2016). Market dynamics and innovation management on performance in SMEs: Multi-agent simulation approach. *Procedia Computer Science, 91,* 707- 714.

Niewöhner, N., Asmar, L., Wortmann, F., Röltgen, D., Kühn A., & Dumitrescu, R. (2019). Design fields of agile innovation management in small and medium sized enterprises. *Procedia CIRP, 84,* 826-831.

Oleśków, J., Stachowiak, A., Batz, A., & Fertsch, F. (2017). The level of innovation in SMEs, the determinants of innovation and their contribution to development of value chains. *Procedia Manufacturing, 11,* 2203-2210.

Bagheri, M., Mitchelmore, S., Bamiatzi, V., & Nikolopoulos, K. (2019). Internationalization orientation in SMEs: The mediating role of technological innovation. *Journal of International Management, 25*(1), 121-139.

Kamalian, A., Rashki, M, Hemmat, Z, & Jolfaie, S. (2015) Cooperation networks and innovation performance of Small and Medium-Sized Enterprises (SMEs). *International Journal of Management, Accounting and Economics, 2,* 233-243.

Rahab, S. (2015). Innovativeness model of small and medium enterprises based on market orientation and learning orientation: Testing moderating effect of business operation mode. *Procedia Economics and Finance, 4,* 97-109.

Rifat, K., Meltem, Y., & Gultekin, A. (2016). *Is the relationship between innovation performance and knowledge management contingent on environmental dynamism and learning capability?* Evidence from a turbulent market, Business Research, 2.

Xing, X., Wang, J., & Tou, L. (2019). The relationship between green organization identity and corporate environmental performance: The mediating role of sustainability exploration and exploitation innovation. *International Journal of Environmental Research and Public Health Public Health, 16,* 6.

Xin, Z., & Sun, B. (2016). The influence of Chinese environmental regulation on corporation innovation and competitiveness. *Journal of Cleaner Production, 112,* 1528-1536.

Maletič, M., Maletič, D., Dahlgaard, J., Dahlgaard-Park, S., & Gomiscek, B. (2015). Effect of sustainability-oriented innovation practices on the overall organizational performance: An empirical examination. *Total Quality Management & Business Excellence.*

Conesa, I., Soto-Acosta, P., & Palacios, M. (2016). Corporate social responsibility and its effect on innovation and firm performance: An empirical research in SMEs. *Journal of Cleaner Production.*

Pittawat, U., & Alisara, J. (2017). Enhancing the innovation of small and medium enterprises in food manufacturing through Triple Helix Agents. *Kasetsart Journal of Social Sciences.*

Apanasovich, N., Alcalde, H., & Parrilli, M. (2016). The impact of business innovation modes on SME innovation performance in post-Soviet transition economies: The case of Belarus. *Technovation, 57–58,* 30-40

Castela, B., Ferreira, F., Ferreira, J., & Marques, C. (2018). Assessing the innovation capability of small- and medium-sized enterprises using a non-parametric and integrative approach. *Management Decision, 56*(6), 1365-1383

De Martino, M., & Magnotti, F. (2018). The innovation capacity of small food firms in Italy. *European Journal of Innovation Management, 21*(3), 362-383.

Kritsadee, P., Sanguan, L., & Somnuk, A. (2017). Factor affecting innovativeness of small and medium enterprises in the five southern border provinces. *Kasetsart Journal of Social Sciences, 38*(3), 204-211.

Kim, M., Park, J., & Paik, J. (2018). Factors influencing innovation capability of small and medium- sized enterprises in Korean manufacturing sector: facilitators, barriers and moderators. *International Journal of Technology Management, 76*(3-4).

Guo, Y., Wang, L., Wang, M., & Zhang, X. (2019). The mediating role of environmental innovation on knowledge acquisition and corporate performance relationship—A study of SMEs in China. *Sustainability, 11.*

Guzman, G., Garza-Reyes, J., Pinzón-Castro, S., & Kumar, V. (2018). Innovation capabilities and performance: Are they truly linked in SMEs? *International Journal of Innovation Science.*

Giulia, B. (2016). A support system to manage product and process changes in manufacturing. *IFAC-Papers Online, 49*(12), 1080-1085.

Gasparetto, W., Egger, G., Giusti, A., Rauch, E., Riedl, M, & Matt, D. (2018) Intelligent workpiece carrier for distributed data collection and control in manufacturing environments. *Procedia Manufacturing, 24,* 190-195.

Suthep, B., & Chiwapon, N. (2019). A fuzzy logic combined with lp model for performance evaluation to distribute purchase orders in cluster manufacturing. *Procedia Manufacturing, 30,* 19-25.

Singh, S., Udoncy, E., & Musa, S. (2016). Development of sustainable manufacturing performance evaluation expert system for small and medium enterprises. *Procedia CIRP, 40,* 608-613.

Bilal, M., Sanin, C., & Szczerbicki, E. (2019). Smart Virtual Product Development (SVPD) to enhance product manufacturing in industry 4.0. *Procedia Computer Science, 159,* 2232-2239.

Reza, H., Ray, Z., & Xun, W. (2018). A survey study on industry 4.0 for New Zealand manufacturing. *Procedia Manufacturing, 26,* 49-57.

Torn, I., & Vaneker, T. (2019). Mass personalization with industry 4.0 by SMEs: A concept for collaborative networks. *Procedia Manufacturing, 28,* 135-141.

Silva, J., Mojica Herazo, J., Rojas, R., Pineda, R., Morgado, R., & Varela, N. (2019). Early warning method for the commodity prices based on artificial neural networks: SMEs case. *Procedia Computer Science, 151,* 1243-1248.

Constantinescu, C., Francalanza, E., & Matarazzo, D. (2015). Towards knowledge capturing and innovative human-system interface in an open-source factory modelling and simulation environment. *Procedia CIRP, 33,* 23-28.

Zygiaris, S. (2017). Innovation Chain Intelligence (ICI). *International Journal of Management Technology, 4*(2), 13-21.

Sivalogathasan, V. (2016). How to improve innovation capability? : The role of intellectual capital and organizational motivation of manufacturing firms in Sri Lanka. *Journal of Business Management.*

Aulawi, H. (2018). Improving innovation capability trough creativity and knowledge sharing behavior. *IOP Conference Series: Materials Science and Engineering,* 434.

Pedron, C., Picoto, W., Colaco, M., & Araújo, C. (2018). CRM system: The role of dynamic capabilities in creating innovation capability. *Brazilian Business Review, 15,* 494-511.

Ossi, Y., Jukka, S., Porras, J., & Vesa, H. (2019). Innovation capabilities as a mediator between big data and business model. *Journal of Enterprise Transformation.*

Adamides, E., & Karacapilidis, N. (2020). Information technology for supporting the development and maintenance of open innovation capabilities. *Journal of Innovation & Knowledge, 5*(1), 29-38.

Álvarez-Castañón, L., López-Lira, A., & Estrada-Domínguez, J. (2019). Open innovation as evolutionary path to increase SMEs' capabilities. *Revista Espacios, 40,* 20.

Bullon, J., González, A., Hernández, A., & Queiruga, A. (2017). Manufacturing processes in the textile industry. expert systems for fabrics production. *Advances in Distributed Computing And Artificial Intelligence Journal, 6,* 41-50.

Santiago-Santiago, K., Laureano-Cruces, A., Antuñano-Barranco, A., Domínguez-Pérez, O., & Sarmiento-Bustos, O. (2015). An expert system to improve the functioning of the clothing industry. *International Journal of Clothing Science and Technology, 27,* 1.

Shih-Wen, H., Chu-Hsuan, L., Rong-Qi, C. (2016). An intelligent system for fashion colour prediction based on fuzzy C-means and gray theory. *Color Research & Application, 42,* 2.

Vinod, S., & Poonkuzhali, S. (2018). Improvising the sales of garments by forecasting market trends using data mining techniques. *International Journal of Pure and Applied Mathematics, 119*(7), 797-805.

Sánchez, M., Aguilar, J., Cordero, J., Valdiviezo-Díaz, P., Barba-Guamán ,L., & Chamba-Eras, L. (2016). Cloud computing in smart educational environments: Application in learning analytics as service. New Advances in Information Systems and Technologies. *Advances in Intelligent Systems and Computing, 444,* 993-1002.

Aguilar, J. (2013). Different dynamic causal relationship approaches for cognitive maps. *Applied Soft Computing , 13,* 271-282.

Aguilar, J., Cerrada, M., Mousalli, G., Rivas, F., & Hidrobo, F. (2005). A multiagent model for intelligent distributed control systems. *Lecture Notes in Computer Science, 3681,* 191-197.

Aguilar ,J. (2001). A general ant colony model to solve combinatorial optimization problems. *Revista Colombiana de Computación, 2,* 7-18.

Aguilar, J., Salazar, C., Velasco, H., Monsalve-Pulido, J., & Montoya, E. (2020). Comparison and evaluation of different methods for the feature extraction from educational contents. *Computation, 8.*

Araujo, M., Aguilar, J., & Aponte, H. (2003). Fault detection system in gas lift well based on artificial immune system. *In Proceedings International Joint Conference on Neural Networks, 3,* 1673-1677.

Packianather, M., Davies, A., Harraden, S., Soman, S., & White, J. (2017). Data mining techniques applied to a manufacturing SME. *Procedia CIRP, 62,* 123–128.

Zimmerling, C., Dörr, D., Henning, F., & Kärger, L. (2019). A machine learning assisted approach for textile formability assessment and design improvement of composite components. *Composites Part A: Applied Science and Manufacturing, 124.*

Klement, N., Silva, C., & Gibaru, O. (2017). Solving a discrete lot sizing and scheduling problem with unrelated parallel machines and sequence dependent setup using a generic decision support tool. *In Proceedings IFIP International Conference Advances in Production Management Systems, 513,* 459–466.

Pfrommer, J., Zimmerling, C., Liu, J, Kärger, L., Henning, F., & Beyere, J. (2018). Optimisation of manufacturing process parameters using deep neural networks as surrogate models. *Procedia CIRP, 72,* 426-431.

Murphy, R., Newell, A., Hargaden, V., & Papakostas, N. (2019). Machine learning technologies for order flowtime estimation in manufacturing systems. *Procedia CIRP, 81,* 701-706.

Zhou, Y., Wang, M., Chen, R., Wang, Y., & Wang, J. (2019). Research on application of SME manufacturing cloud platform based on micro service architecture. *Procedia CIRP, 83,* 596-600.

Weldeslasie, D., & Ahmed, D. (2019). The application of expert systems in manufacturing sector. *International Journal of Innovative Technology and Exploring Engineering, 9*(1), 4744-4749.

Arjmandi, M., Ramezani, M., Giordano, M., & Schmid, S. (2017). Finite element modelling of sliding wear in three-dimensional woven textiles. *Tribology International, 115,* 452-460.

Siebert, F. (2019). Research, development and innovation: The challenges of Chile and Italy amid the current technological irruption.

Interaction Design Foundation. (N.D). Literature topics design thinking. What is design thinking?

Ries, E. (2008). The lean startup method. Deusto, Barcelona.

Carayannis, E.G., & Campbell, D.F. (2009). Mode 3 and Quadruple Helix: Toward a 21st century fractal innovation ecosystem. *International journal of technology management, 46*(3-4), 201-234.

Appendix B

Autonomous cycles of data analysis tasks for innovation processes in MSMEs

The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/2210-8327.htm

Autonomous cycles of data analysis tasks for innovation processes in MSMEs

Ana Gutiérrez
GIA, Universidad Francisco de Paula Santander, Cúcuta, Colombia; GIDITICS, Universidad EAFIT, Medellín, Colombia and Tepuy R+D Group, Artificial Intelligence Software Development, Mérida, Venezuela
Jose Aguilar
GIDITICS, Universidad EAFIT, Medellín, Colombia; CEMISID, Universidad de Los Andes, Núcleo La Hechicera, Mérida, Venezuela; Dpto. de Automática, Universidad de Alcalá, Alcalá de Henares, Spain and Tepuy R+D Group, Artificial Intelligence Software Development, Mérida, Venezuela
Ana Ortega
GIDITICS, Universidad EAFIT, Medellín, Colombia, and
Edwin Montoya
GIDITICS, Universidad EAFIT, Medellín, Colombia and Tepuy R+D Group, Artificial Intelligence Software Development, Mérida, Venezuela

Received 24 February 2022
Revised 10 April 2022
30 April 2022
Accepted 5 May 2022

Abstract

Purpose – The authors propose the concept of "Autonomic Cycle for innovation processes," which defines a set of tasks of data analysis, whose objective is to improve the innovation process in micro-, small and medium-sized enterprises (MSMEs).

Design/methodology/approach – The authors design autonomic cycles where each data analysis task interacts with each other and has different roles: some of them must observe the innovation process, others must analyze and interpret what happens in it, and finally, others make decisions in order to improve the innovation process.

Findings – In this article, the authors identify three innovation sub-processes which can be applied to autonomic cycles, which allow interoperating the actors of innovation processes (data, people, things and services). These autonomic cycles define an innovation problem, specify innovation requirements, and finally, evaluate the results of the innovation process, respectively. Finally, the authors instance/apply the autonomic cycle of data analysis tasks to determine the innovation problem in the textile industry.

Research limitations/implications – It is necessary to implement all autonomous cycles of data analysis tasks (ACODATs) in a real scenario to verify their functionalities. Also, it is important to determine the most important knowledge models required in the ACODAT for the definition of the innovation problem. Once determined this, it is necessary to define the relevant everything mining techniques required for their implementations, such as service and process mining tasks.

Practical implications – ACODAT for the definition of the innovation problem is essential in a process innovation because it allows the organization to identify opportunities for improvement.

Ana Gissel Gutiérrez Buitrago is supported by a PhD grant financed by Universidad EAFIT. All the authors would like to thank the "Vicerrectoría de Descubrimiento y Creación" of Universidad EAFIT, for their support on this research.

Applied Computing and Informatics
Emerald Publishing Limited
e-ISSN: 2210-8327
p-ISSN: 2634-1964
DOI 10.1108/ACI-02-2022-0048

Originality/value – The main contributions of this work are: For an innovation process is specified its ACODATs in order to manage it. A multidimensional data model for the management of an innovation process is defined, which stores the required information of the organization and of the context. The ACODAT for the definition of the innovation problem is detailed and instanced in the textile industry. The Artificial Intelligence (AI) techniques required for the ACODAT for the innovation problem definition are specified, in order to obtain the knowledge models (prediction and diagnosis) for the management of the innovation process for MSMEs of the textile industry.

Keywords Innovation process, Autonomic computing, Artificial intelligence, Data analytics

Paper type Case study

1. Introduction

Micro-, small- and medium-sized enterprises (MSMEs) have limited resources, and thus, they must search for efficient ways to do more with less [1, 2], especially in the quarantine economy [3, 4] in light of coronavirus disease 2019 (COVID-19) [5, 6]. Particularly, MSMEs need to innovate and improve their offer of goods, products and services, to respond to the changing needs of the market. Innovation has become the means that allows an MSME to transform and continue to grow to stay in the market, taking advantage of each of the resources available in the organization, human, technological and financial. Several studies have concluded that investment in innovation and technology has an impact on the development of organizations to be more competitive, which leads many times to the introduction of new products and processes [7, 8]. In turn, the return on investment will be reflected in productivity indicators, in good operation and profitability of the organization.

On the other hand, information is becoming more relevant every day for companies to make decisions. Organizations not only need to collect data but also find the right way to analyze it to devise daily actions based on statistics and trends. However, companies currently lack the capacity to use big data and data analytics [9]. Therefore, companies must start using all available data sources, and be able to make the most of data to support decision-making in their organizations. Especially, it is necessary to understand and analyze the different sources of information that will improve the innovation processes with the use of data analytics tasks, to respond to the different phases of them.

Given the importance of the innovation in MSMEs, and the current opportunities that exist to exploit data from the organizations and their contexts can be defined strategies based on data to build data-driven models to guide the innovation processes. One of these strategies is the utilization of the concept of “autonomous cycles of data analysis tasks” (ACODATs) defined in previous works [10–12], which allow generating knowledge models useful for the management of the innovation processes using different data sources. An ACODAT is composed by a set of data analysis tasks to reach a goal for a given problem, where each task has a given role [13–15]: observe the studied system, analyze it and make decisions to improve it. In this way, there are interactions and synergies between the data analysis tasks, to generate the required knowledge with the goal of improving the process under study.

In this paper, we propose several ACODATs for the management of the innovation processes in an MSME. Likewise, in the paper is proposed the specification in detail of the autonomic cycle for the innovation problem definition sub-process, and its application in the textile industry. For the development of the ACODATs, the MetodologIa para el Desarrollo de Aplicaciones de Minería de Datos basados en el aNálisis Organizacional (MIDANO) [16–18] methodology was used, which allows the development of data analytics applications, and especially, the development of ACODATs. The main contributions of this work are:

(1) The specification of ACODATs for the management of innovation processes.

(2) The definition of a multidimensional data model, which stores the required information of the organization and the context for the ACODATs.

(3) The detailed description of the ACODAT for the definition of the innovation problem, which is instanced in the textile industry.

(4) The characterization of the AI techniques required for the ACODAT for the innovation problem definition, in order to obtain the knowledge models (prediction and diagnosis) for the management of the innovation processes for MSMEs of the textile industry.

This work is organized as follows. Section 2 presents the related works. In Section 3, the theoretical framework is presented, specifically, ACODAT, MIDANO and the innovation model used in this work. Section 4 introduces the autonomic cycles proposed, the description of their tasks and their multidimensional data model, using the MIDANO methodology. Section 6 details the case study of the textile industry, and the application of the autonomic cycle for the definition of innovation problems. Finally, the conclusions and future works are presented.

2. Related works

In this section, we present the main recent papers related to our approach, which are the definition of schemes for the automation of innovation processes or the utilization of autonomic cycles in the automation of industrial processes (Industry 4.0).

Ossi *et al.* [19] presented a conceptual framework based on big data and business models to exploit the innovation capabilities. The framework adopted the business canvas model. This framework helps in concentrating on different viewpoints, for example, can create and develop strategies of price based on analytics data. The framework offers ways to organize perspectives for organizational transformation. On the other hand, machine learning (ML) models offer the computational power and functional flexibility required to decipher complex patterns in a high-dimensional data environment [20]. Particularly, in [20] three groups of financial data analysis are identified: (1) portfolio management; (2) financial fraud and distress; and (3) sentiment inference, forecasting and planning.

Kritsadee *et al.* [21] tested a model of factors affecting the innovativeness of small and medium enterprises (SMEs) using the structural equation model (SEM). Data about innovativeness were collected using questionnaires, which were mailed to 283 entrepreneurs. The proposed model determined that learning orientation and proactiveness had direct effects on innovativeness. The analysis addressed the innovation in products, processes, organizational and marketing, and their contribution to the organization's results (e.g. market share, environmental sustainability, profit, etc.). The paper [22] investigated the parameters in the innovation process design that influence the innovation outcomes in the context of smart manufacturing (Industry 4.0), and thus what should be accounted for in the design of innovation processes for smart manufacturing. The research is based on empirical evidence from 18 manufacturing companies and suppliers of manufacturing technology. Finally, the authors of [23] present a systematic literature review about how smart systems have been used to improve the innovation capacities in MSMEs. The results show that there is not an established body of knowledge about how to improve the innovation process based on smart systems.

Sanchez *et al.* [15] defined three autonomic cycles that allow interoperating the actors of manufacturing processes (data, people, things and services). Particularly, they defined a framework for the integration of autonomous processes based on cooperation, collaboration and coordination mechanisms. The framework is composed of three ACODATs that allow the self-configuration, self-optimization and self-healing of the manufacturing process. They implement one of these ACODATs, for the self-supervision of the coordination process mixing it with the theory of multi-agent systems [24]. This ACODAT is implemented and tested using an experimental tool that replays a production process event log, to detect failures and invoke the ACODAT for self-healing when needed. Qin *et al.* [25] proposed a multi-layered framework of manufacturing for Industry 4.0. One of the levels, the intelligence

layer, applies different data analytic tasks to discover useful information from data to improve the manufacturing process. Thus, the intelligence layer creates a knowledge base that serves as a support for the planning and decision-making processes.

Besides, the paper [26] reveals that knowledge management for sustainability research has relied on nine foundational clusters (i.e. informed sustainability practice, social network, firm performance, knowledge sharing culture, green innovation, sustainability assessment framework, global warming, knowledge management and innovative performance) to generate new knowledge. Also, they determine that the method of creating, communicating, disseminating and exploiting shared knowledge is instrumental for firms adopting business practices to enhance firm performance.

The previous studies do not define frameworks and systems for the management of the innovation processes for MSMEs based on the ACODAT concept, neither do they clarify the application of data analytic to improve the innovation capabilities in an organization. These are the main differences in our approach with respect to previous works. On the other hand, the ideas proposed in this work could be used in other areas of an organization, including environmental social governance (ESG) and total quality management (TQM) [27].

3. Conceptual framework

3.1 ACODAT

This research follows the ACODAT concept, which is based on the idea proposed by IBM in 2001 [28]. The ACODAT concept was proposed in [10–12, 29] and has been used in telecommunication [30], education, especially in smart classrooms [11, 12], Industry 4.0 [13–15] and smart cities [31], among other domains. It is based on the autonomic computing paradigm [32], with the purpose of endowing autonomic properties to systems based on a smart control loop.

The main objective of an ACODAT is to extract useful knowledge from data to make decisions [11, 12]. The set of data analysis tasks must be performed together, in order to achieve the objective in the process supervised. The tasks interact with each other and have different roles in the cycle, which are: observing the process, analyzing and interpreting what happens in it and making decisions to reach the objective for which the cycle was designed. This integration of tasks in a closed loop allows solving complex problems. The detailed description of the roles of each task is [11, 12]:

Monitoring: Tasks to observe the supervised system. They must capture data and information about the behavior of the system. Besides, they are responsible for the preparation of the data for the next step (preprocessing, selection of the relevant features, etc.).

Analysis: Tasks to interpret, understand and diagnose what is happening in the monitored system. These tasks allow building knowledge models about the dynamics observed, in order to know what is happening in the system.

Decision-making: Tasks to define and implement the necessary actions based on the previous analyses, in order to improve the supervised system. These tasks impact the dynamics of the system, and their effects are again evaluated in the monitoring and analysis steps, restarting a new iteration of the cycle.

In general, an ACODAT requires:

(1) A multidimensional data model that represents the data collected from the different sources, in order to characterize the behavior of the context, which will be used by the different data analysis tasks.

(2) A unique platform to integrate the different technological tools required by the data analysis tasks to carry out data mining, semantic mining and linked data, among others.

This concept has been successfully proven in different fields, but ACODAT has not been applied in innovation processes.

3.2 MIDANO

MIDANO is a methodology for the development of data analytics-based applications [16, 18], which is made up of three phases:

Phase 1 – Identification of data sources for the extraction of knowledge of an organization: This phase carries out a knowledge engineering process-oriented to organizations/companies. The main objective of this phase is to know the organization, its processes and its experts, among other aspects, to define the objective of the application of data analysis in the organization. Also, it defines the autonomic cycles and their data analysis tasks.

Phase 2 – Preparation of data: To apply data analysis to a specific problem, it is necessary to have data associated with the problem. This involves performing different operations with the data, with the purpose of preparing them. This process is based on the paradigm ETL: extraction of data from the sources, data transformation and loading of the data in a data warehouse. During this phase are described all the variables of interest and carried out the data processing process (for example: dependency analysis among variables, normalizations, etc.). Also, this phase designs the multidimensional data model of the autonomic cycles, which is the structure of the data warehouse. Finally, it carries out a feature engineering process that consists on transform raw data into features. A feature engineering process includes the tasks of extraction, generation, fusion and selection of variables for the construction of the knowledge models.

Phase 3 – Development of the autonomous cycle: In this phase, the data analysis tasks are implemented, which are going to generate the required knowledge models (e.g. predictive and descriptive models). This stage culminates with the implementation of a prototype of the autonomic cycle. This phase can use existing data mining methodologies for the development of the data analysis tasks. In addition, during this phase, experiments are carried out to validate the knowledge models generated.

3.3 Proposed model of innovation processes

The innovation process is a structured strategy that ensures that the innovation team idealizes an innovation and executes it until its successful implementation. In this section, we explain the innovation process model defined in [23]. According to [23], an innovation process has four sub-processes: problem analysis, ideation, experimentation and commercialization. Each phase (sub-process) is described below.

(1) Problem analysis: The problem must be identified and defined.
- Definition of the problem: This step must indicate and define the problem.
- Specification of needs: It defines a list of requirements necessary to solve it.

(2) Ideation: It defines the concepts to develop.
- Generation of many ideas: In this step are generated ideas. The amount here matters. The more, the better. It can use the technique of brainstorming
- Ideas evaluation: It is the process of comparing and contrasting ideas related to the new product, to select the most promising.
- Selection of the best idea: The idea that best solves the problem is selected.

(3) Experimentation: In this step is generated a version, although not be exact to the initially proposed product.

- Prototype: It is the development of an initial product, which allows deciding if it is feasible.
- Test: The main objective is to validate the creative process.
- Escalation: It transforms a concept (prototype) in a commercial product.

(4) Commercialization: It is the process of launching new products or services to the market.

- Launching: It is oriented to publicize the innovative product and its results.
- Results measurement: It defines the metric to measure the results of the marketing process.
- Learning cycle: The market will give feedback to know if the idea must be changed, optimized or persevere with it.
- Internal diffusion: It is the communication between the workers. The objective is the utilization of innovation as a positive reinforcement to motivate the organization.

4. Application of MIDANO for the definition of autonomic cycles for an innovation process

In this section is analyzed an innovation process using the MIDANO methodology, in order to define the sub-processes where the autonomic cycles must be defined.

4.1 Sub-processes of an innovation process

An innovation process has different sub-processes, which must be prioritized according to if data analysis tasks can be used. There are 12 sub-processes defined in an innovation process, which are listed in Table 1.

4.2 Prioritization

The criteria to be considered to evaluate the relevance of the sub-processes were defined according to their importance for an innovation process (especially, for a textile organization) and the possibility to carry out data analysis tasks. Thus, these values determine the level of

Processes	Sub-processes	Abbreviation
Problem analysis	Definition of the problem	DDP
	Specification of needs	EDN
Ideation	Generation of ideas	GDI
	Ideas evaluation	EDI
	Selection of the best	SDM
Experimentation	Prototyped	PRO
	Pilot test	PPI
	Escalation	ESC
Commercialization	Launch	LAN
	Results measurement	MDR
	Learning cycle	CDA
	Internal dissemination	DIN

Table 1.
Sub-processes of the innovation process

importance of the sub-processes. For example, a process that is not important has a weight of 1 and a process very important has a weight of 5.

The case study is in the textile sector because it is one of the industrial sectors where MSMEs require more continuous innovation processes, to enable them to be competitive over time [23]. Likewise, it is the industrial sector of interest for the context where the project is developed, for which data are available to carry out data analysis tasks to improve it.

For the construction of the prioritization table, 10 experts from the fashion innovation sector and research professors were consulted, who participated by qualifying each of the criteria. In the final result, each of the answers provided by the experts was averaged. Results are shown in Table 2.

From the previous table, the sub-processes "Problem Definition", "Specification of Needs" and "Measurement of Results" were prioritized. The sub-process "Definition of the Innovation Problem" was the one that had the highest evaluation among the sub-processes because, in most of the criteria evaluated by each of the experts, its rating was equal to or greater than 4. It has a very good rating in each group of criteria: about the possibility to apply data analysis tasks in the process, how it impacts the innovation process and its interest in the textile industry. Particularly, in some criteria about its importance in the innovation process, it has the highest score (its impact in the innovation process and in the generation of new products and services, with a rating of 5).

4.3 Analysis of the strategic objectives to be achieved with these sub-processes using autonomic cycles

For the prioritized sub-processes in Table 2, it is required to characterize the current situation in each one. Table 8 in section "Supplementary Material" contains the actors involved in the sub-process, the data sources and activities that are used and the obtained results (goal to be reached). These results now must be reached using data analytic tasks.

5. Definition of the autonomic cycles

This section presents the ACODATs of the prioritized sub-processes, in order to enable autonomic coordination in the innovation processes (ACIP-000, see Figure 1), but particularly, it describes the design of the sub-process of the definition of the innovation problem.

The goal of ACIP-000 is the self-management of the innovation processes. In order to reach this goal, we propose three ACODATs:

ACIP-001 (Innovation Problem Definition): This cycle is responsible for obtaining useful information for the definition of the innovation problem. The goal of this autonomic cycle is the definition of the innovation problem based on the information of the organization and context.

ACIP-002 (Specification of Needs): This cycle is responsible for obtaining the requirements to be covered by the innovation process. The goal of this autonomic cycle is the identification and characterization of the requirements of the innovation problem.

ACIP-003 (Result Measurement): This cycle is responsible for assessing the quality of the results obtained during the innovation process. The goal of this autonomic cycle is the definition of the strategies and metrics to evaluate the results of the innovation process, and the evaluation of the results to determine the quality of the innovation process.

We have proposed three ACODATs according to the sub-processes prioritized in section 4.2 (ACIP-001, ACIP-002, ACIP-003). This prioritization was made according to the relevance of the sub-processes for the innovation processes of an organization and the possibility of automating them using data.

However, it is important to mention that there are other sub-processes in the model of innovation processes defined in section 3.3. They could be specified in the future using

WEIGHT	CRITERIA TO EVALUATE	PROCESSES											
		PROBLEM ANALYSIS		IDEATION PROCESS			EXPERIMENTATION			COMMERCIALIZATION			
		DDP	EDN	GDI	EDI	SDM	PRO	PPI	ESC	LAN	MDR	CDA	DIN
	Relevance in the innovation process												
5	How quality impacts your innovation processes	5	5	4	4	5	4	4	4	4	4	4	4
5	How much depends on the other organizational processes	4	4	4	4	4	4	4	4	4	4	3	4
5	The complexity of the personnel and infrastructure required	4	4	4	4	4	5	5	4	4	4	4	4
	How the generation of new products or services impacts	5	5	4	4	5	4	4	4	4	4	4	4
5	How it impacts the business processes and the organization	4	4	3	4	4	4	3	4	4	4	4	4
5	How it impacts on marketing strategies	4	4	4	4	4	4	4	4	5	5	4	4
5	How much it encourages creativity	4	4	5	4	5	4	4	4	4	3	4	4
	Relevance for a textile organization												
	How much it impacts the organization	4	4	3	4	4	4	3	4	4	4	4	4
5	The level of difficulty to be performed in practice	4	4	3	4	4	4	4	5	4	3	4	3
5	How much help in the consolidation of an intelligent Textile Company (Smart Textile Factory)	4	4	4	4	4	4	4	4	3	4	4	4
	Relevance to do Data Analytics tasks												
5	How many internal or external sources of information exist: databases, Excel sheets, reports	4	4	4	4	3	3	3	3	3	4	3	3
5	What level of access do you have to the Information	4	4	4	4	4	4	3	3	3	4	3	3
5	The level of use of computational tools (word, excel, power point, etc.)	3	4	4	4	3	4	4	4	4	5	4	4
5	The level of use of decision-making software (DSS) - (CRM-ERP-MRP-CIM-COGNOS-informaweberp)	4	4	3	3	3	3	4	4	4	4	4	3
5	The frequency of information collection in this phase of the process	4	4	4	3	3	3	4	4	4	5	4	3
5	Total unweighted	62	61	57	57	58	58	57	58	57	61	57	53
5	Weighted total	312	303	287	284	290	291	285	292	283	306	286	264

Table 2.
The prioritized sub-processes

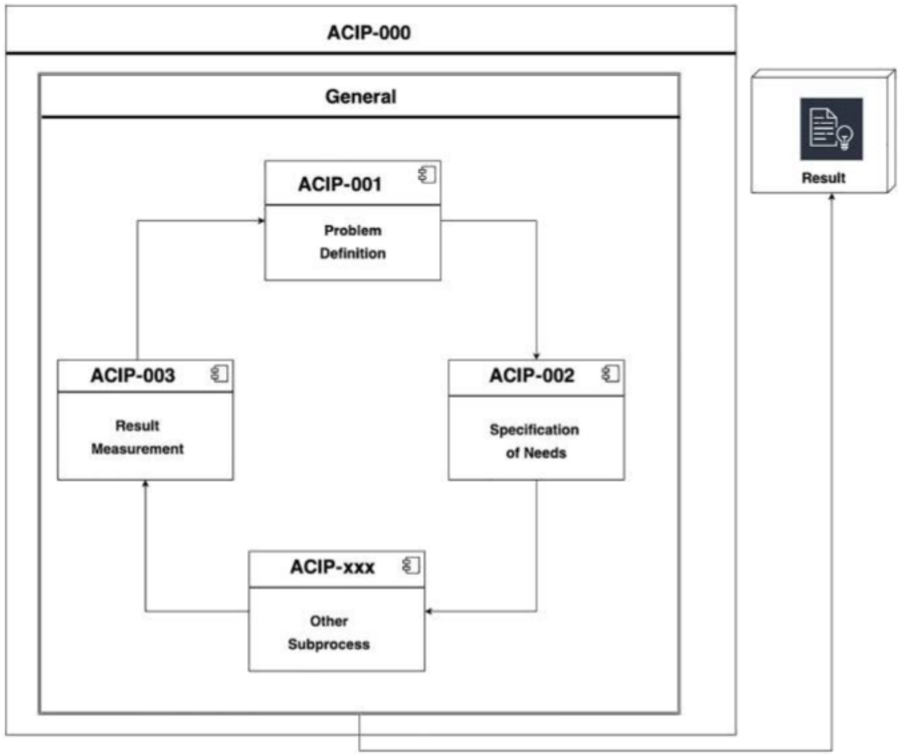

Figure 1. ACIP-000: Prioritized autonomic cycles of an innovation process

ACODATs to automate them as well. Thus, ACIP-xxx refers to ACODATs for the other innovation sub-processes, such as generation of many ideas, ideas evaluation, selection of the best idea, among others.

Finally, the alerts module is an information system on the execution status of an innovation process (started, executed, finished), and additionally, it would inform about which of the sub-processes would be running.

In this article, we detail the ACIP-001, which was the one that obtained the highest evaluation in the prioritized processes.

5.1 Specification of the autonomic cycles for the "definition of the problem"

The Autonomous Cycle for the Innovation Problem Definition (ACIP-001 – Problem Definition) has as its main objective the characterization of the innovation problem, i.e. the statement of the problem. In general, this autonomic cycle is defined by a set of data analysis tasks, which use everything mining techniques to get useful information to create the statement of the innovation problem. We use the 5Ws model to define this cycle because it allows defining what the problem is and not the solution (see Figure 2). The 5Ws model was established by the Greek rhetorician Hermagoras of Tendon, from where it has evolved [33]. In the 5W model, each question must obtain an answer based on specific data.

Table 3 shows the general description of each task of this autonomic cycle.

Now, we describe each task.

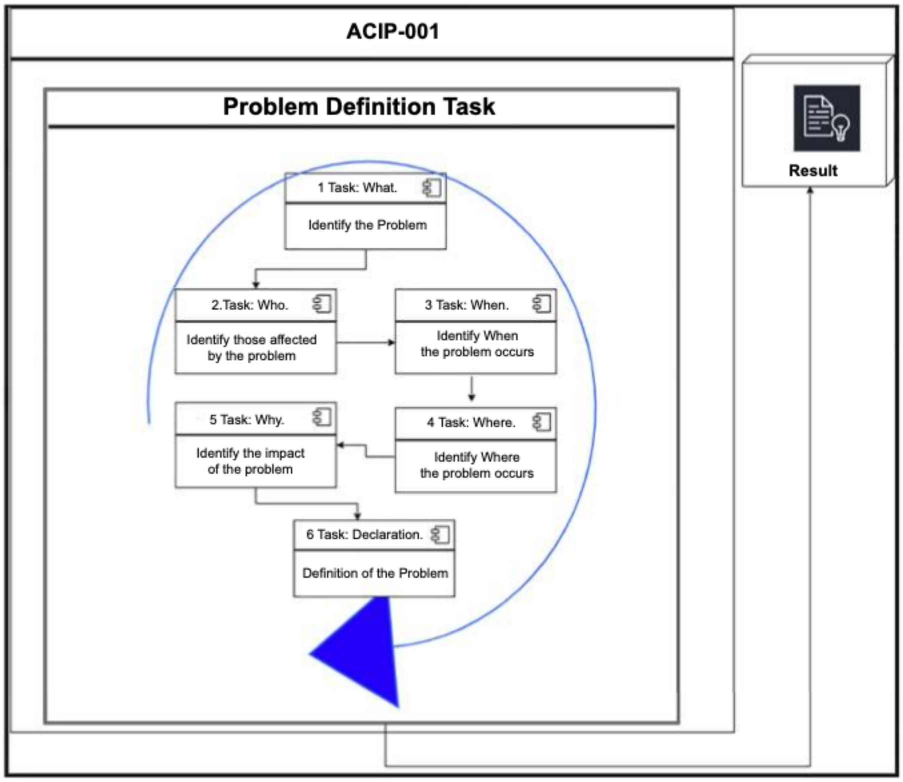

Figure 2.
Structure of the ACIP-001

Table 3.
Description of the tasks of ACIP-001

Task name	Knowledge models	Data sources
1. What: Identify the problem	Descriptive model Detection model	Market studies (customer opinions, satisfaction surveys, etc.), internal databases (CRM[1], PQRS[2]) Social networks (Instagram, Facebook, etc.)
2. Who: Identify those affected by the problem	Descriptive model	Market studies (customer opinions, satisfaction surveys, etc.), internal databases (CRM[1], PQRS[2]) Social networks (Instagram, Facebook, etc.)
3. When: Identify when the problem occurs	Detection model Predictive model	Market studies (customer opinions, satisfaction surveys, etc.), internal databases (CRM[1], PQRS[2]) Social networks (Instagram, Facebook, etc.)
4. Where: Identify where the problem occurs	Diagnostic model Predictive model	Market studies (customer opinions, satisfaction surveys, etc.), internal databases (CRM[1], PQRS[2]) Social networks (Instagram, Facebook, etc.)
5. Why: Identify the impact of the problem	Diagnostic model Predictive model	Market studies (customer opinions, satisfaction surveys, etc.), internal databases (CRM[1], PQRS[2]) Social networks (Instagram, Facebook, etc.)
6. Declaration: Definition of the problem	Natural Language Processing (NLP)	

Note(s): [1]CRM: Customer Relationship Management
[2]PQRS System (Requests, Complaints, Claims and Suggestions)

(1) *Task 1. What: Identify the problem*: The first step identifies the problem through the data obtained. Some examples of data sources can be quality problems, customer complaints or derived from competitive surveillance activities. Its objective is to determine the occurrence of an innovation problem (i.e. it is necessary to create an original solution). This task uses detection and descriptive models to identify the problem.

(2) *Task 2. Who: Identify those affected by the problem*: This task identifies who are affected by the problem (e.g. specific groups, organizations, customers). This task uses descriptive models.

(3) *Task 3. When: Identify when the problem occurs:* This task identifies when the problem occurs or will occur, for which it can use detection or prediction models.

(4) *Task 4. Where: Identify where the problem occurs:* This task identifies where the problem is occurring, for which it uses diagnosis models.

(5) *Task 5. Why: Identify the impact of the problem:* This task identifies the importance of the problem, for this, it seeks to answer questions such as: What impact does it have on the business? What impact does it have on all stakeholders (i.e. employees, suppliers and customers)?

(6) *Task 6. Declaration: Definition of the problem*: This task aims to define the problem statement. For this, it uses NLP techniques to define the narrative.

Finally, the results module is a dashboard to report the execution status of this ACODAT, in particular, the results of its tasks. For example, when task 1 finishes, then it shows the information of the negative twitters; or when task 6 finishes, then it reports the problems that have been defined.

5.2 Multidimensional data model

The multidimensional data model for the previous ACODATs is defined in this section.

The model in Figure 3 includes different data sources, from market studies (e.g. customer opinions, satisfaction surveys), organizational databases (e.g. CRM, PQRS), until social networks (e.g. Instagram, Facebook). Data from each source are included in a different dimension in the data model, according to its characteristics. The main dimensions are the following:

Customers: It stores customer data such as age, gender, marital status, occupation, income, level of education, nationality, direction, country, department, municipality, neighborhood and stratum.

Market study: Stores general market study information, such as the objective, hypothesis, kind of investigation, type of analysis and conclusions. Also, it is linked to other dimensions like:

(1) Product satisfaction: It stores the satisfaction rating data of a product resulting from surveys that answer questions such as, what do you like the most, changes to improve, characteristics of other products that you would like in this product, product comfort, user experience, etc.

(2) Product price: It stores product price sensitivity data such as if you know the product, would you pay more or less to get it? Product units that you would buy taking into account the reduction or increase in price? Money willing to pay, a reasonable price, brand trust, factors that influence the purchase decision and what you like best about the product?

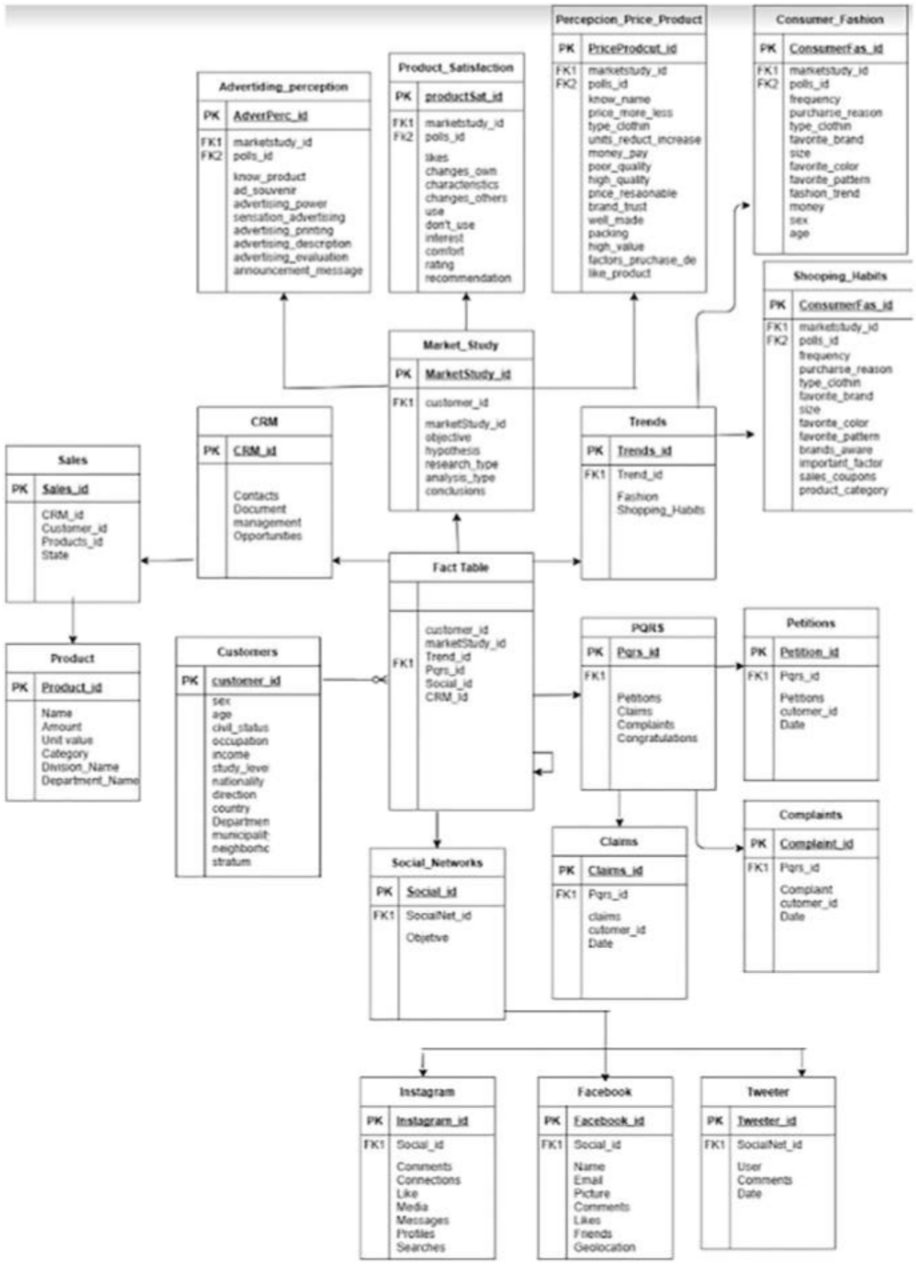

Figure 3.
Multidimensional data model

(3) Advertising perception: It stores data on the perception of advertising, such as product knowledge, recall of the ad, evaluation of the power of advertising, feeling you have when you see an advertisement, the impression that the advertising gives,

how would you evaluate the advertising in comparison with other publications of the competition? and what would be the main message of the advertisement?

Trends: it describes information about current fashion.

(4) Consumer fashion: It stores data on fashion and its consumers, resulting from surveys that answer questions such as frequency of buying new clothes, reasons for shopping, clothing type, favorite brand, favorite color, favorite pattern, trend tracking, gender and age.

Social networks: It describes information about the social networks (Twitter, Instagram, etc.). For example,

(5) Instagram: It stores data of this social network, such as connections (contacts), likes, etc.

The multidimensional data model depicted in Figure 3 includes all the data required by the ACODATs. It describes all the variables of interest, which will be used as data sources to build the knowledge models (descriptive, predictive, among others) defined in each of the tasks of the ACODATs. This will allow having the necessary information to apply the different data analysis techniques to reach the goal of each ACODAT.

6. Case study

This section presents the experimental context for the instantiation of ACIP-001 (Innovation Problem Definition).

6.1 Experimental context

In this case study, we used data from the "Ramara Jeans" store, in Cucuta, Norte de Santander - Colombia. The store is dedicated to the manufacture, sale and marketing of all kinds of jeans, pants, shorts and skirts. Its objective is to provide the best service and quality in the products it offers, becoming a leader in the production of comfortable, versatile garments with competitive prices in the market.

The store currently has social networks on Facebook like Ramara Cúcuta, and Instagram like Ramara Jeans and on WhatsApp a line 313-8092414. It also has a team dedicated to virtual sales of products nationwide to attend to all requests, doubts and questions from its customers. The dataset used in this instantiation is from Instagram.

6.2 Instantiation of the ACIP–001: definition of the problem

At the beginning of the innovation process, it is necessary to define the problem. In this section, we describe how the ACIP–001 is instantiated in this case study.

(1) *First task:* This task can use descriptive and detection models to group and detect potential customer problems according to the client behaviors on the web, customer complaints on social networks, etc. Table 4 shows an example of a log file in an

Store id	Customer id	Age	What problem
Store-1	000110	35	Long waiting or delivery times
Store-1	000111	35	They would not recommend the brand
Store-1	000112	29	Product return
Store-1	000113	41	Low product quality
Store-1	000114	26	Abandonment of the purchase

Table 4. "What": Information generated by the first task

organization, which can be built from a social network (using NLP techniques) or a PQRS database. The last column describes the results of the reported information by the clients.

Also, we can carry out a sentimental analysis to determine the negative sentiments in the social network (maybe due to a problem). For example, we can analyze the client's tweets (see Table 5). If a tweet is negative, it could be a complaint or the presence of a problem. For this task, the priority is to analyze the negative tweets (sentiment = 0) to identify the problem.

For this task, it is necessary to execute an NLP process to detect the problem in the negative tweets, which is composed of the next tasks: tokenize, remove stop words, clean special characters and stemming/lemmatization.

(2) *Second task:* It uses the information collected in the previous step to identify the person who is affected by the problem. In this case, this person could be an online customer face-to-face client consumer, etc. We can use a descriptive model that groups the clients according to the problem, in order to determine the type of clients affected by this problem.

For example, in Figure 4 are shown three different clusters (groups of customers) for three different problems. In this case, one of them are well-differentiated (cluster 1, which has only loyal customers). Cluster 2 (green impulsive customers) has some overlap with cluster 3 (customers by necessity).

(3) *Third task:* This task identifies when occurs the problem, which may occur before the purchase, due to some damage to the garment, or after the purchase. Examples are that the garment is very small or large, that the texture is very bad, etc. In Table 6, the column "when" represents the results of a predictive model about when the problem occurs: (0) before the purchase and (1) after the purchase. Also, we can use a detection model in order to detect in real-time a problem.

(4) *Fourth task*: This task identifies where the problem occurs. In this case, it is very important to identify the context of the problem, for which can be used a diagnosis model. In Table 6, the column "where" shows the results of a predictive model to determine where the problem occurs: (0) according to the customer's perspective or (1) into the organization. Also, it is possible to use a diagnosis model for the same problem.

(5) *Fifth task*: This task identifies the importance of solving the problem. For that, it can diagnose or predict the impact of the problem. In Table 6, the column "impact" shows the results of a predictive model about the impact of the problem. The value (0) is low impact, (1) is medium impact and (3) is high impact.

Table 5. Identify negative tweets

Id	Tweets	Sentiment	Sentiment_ description
1613	@cruuzzy All the products are so beautiful . . .	1	Positive
5048	@radicalj Marvellous - not. How very thwartin. . .	0	Negative
9955	@annajudithk Delivery times are too long:(	0	Negative
1318	The cloth is always nice:-)	1	Positive
6739	@Bett_Homes there are not many products in the store. . .	0	Negative
4179	I love the blue and black ones:)	1	Positive
4084	@MarkBreech Not sure it would be a good thing 4 . . .	1	positive
1798	I loved them, I want another two	1	positive
3892	thank you for your response to my question	1	positive
3027	Hey @Indie_Shell Thanks For Following:) \n\n#. . .	1	positive

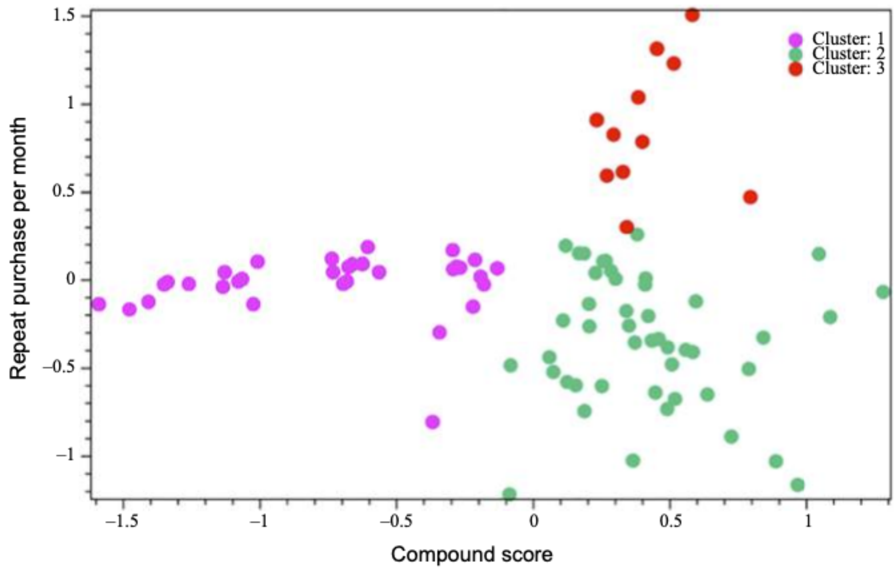

Figure 4.
Who: Clusters to determine customer types

Table 6.
When and where: Predictions generated by the third, four and fifth tasks

Store id	Problem id	Description	When	Where	Impact
Store-1	000001	Long waiting or delivery times	1	1	3
Store-1	000002	They would not recommend the brand	0	0	1
Store-1	000003	Product return	1	0	1
Store-1	000004	Low product quality	1	1	1
Store-1	000005	Abandonment of the purchase	0	0	3

(6) *Sixth task*: It defines the problem taking into account the results of each of the previous tasks. In this task can be used NLP to define the statement of the problem in order to combine the what, who, when, where and why results. Additionally, we can add more information on the context using data from the reviews, tweets, etc. For example, we can use the information of the negative tweets (e.g. the keywords of their texts, determined by metrics such as TF-IDF) [34]. Some examples of statements of a problem, in this case study, are:

- “Long waiting or delivery times” is a “problem with high impact” “after the purchase”
- “Abandonment of the purchase” is a “problem with high impact” “before the purchase”
- “Long waiting or delivery times” because “Delivery times are too long”
- “They would not recommend the brand” is a “problem according to the customer’s perspective”

7. Results discussion

The main result of this work is the definition of different ACODATs for the management of the innovation processes in an organization and the detailed description of the autonomous cycle for the sub-process of innovation problem definition. For this, the data analysis tasks of the cycles were defined and the data sources were identified. Each task builds an appropriate knowledge model using the respective data sources to accomplish its specific objective. For example, in the case study, the first task carried out a sentiment analysis on tweets to identify the problem, and the second task carried out a clustering model to identify the types of users for each problem.

In particular, this autonomous cycle defines the fundamental input for the model of innovation processes proposed in section 3.3: the possible problems that are sources of innovation. Some of these identified problems will later be converted into an innovative product following our model. For example, in the case study, "Long waiting or delivery times" identifies a problem in the final delivery of the product that should lead to innovation in the purchase delivery processes. Another example is "Abandonment of the purchase", which identifies the disinterest shown by customers when they are about to buy a product. This may imply requiring innovation in product presentation/marketing strategies.

Another important result to highlight is the prioritization of sub-processes. To do this, the potentially automatable sub-processes of the innovation model proposed in section 3.3 were first analyzed using the organization and environment data. Subsequently, using the opinion of the experts, it was determined which of them is more relevant (priority) to automate in an initial process of automation of the management of the innovation processes in an organization. For this, the MIDANO methodology was used (see sections 3.2 and 4), which also, allowed defining the ACODATs and designing the autonomous cycle for the first prioritized sub-process (see section 5).

Another result is the definition of the data multidimensional model to be used by the ACODATs. It identifies the set of variables that must be used by the tasks of the ACODATs. With them, the data analysis tasks can build the different knowledge models (predictive, descriptive, etc.) which later are used to reach the goal of each autonomous cycle.

Finally, in the case study is instanced the first autonomous cycle, whose main objective is the identification of problems that potentially will be sources of innovation processes in the organization. In particular, it defines a sentiment analysis task to identify twitters that potentially describe a problem. It then groups those tweets by customer types. It then uses predictive models to determine when and where these problems occur, and their impacts. Finally, it performs a PLN process to formulate the sentences of these problems and potential sources of innovation processes.

It is the first step in demonstrating that it is possible to apply artificial intelligence techniques to improve innovation processes. It is a challenge to implement the rest of the ACODATs, but the preliminary results encourage the continuation of the application of these techniques in the innovation processes in the organizations.

8. Comparison with previous works

In this section, we propose criteria to compare our proposition of autonomic cycles to automate the innovation processes with other works. We define the next criteria:

(1) *Criterion 1:* they automate one of the sub-processes (e.g. definition of the innovation problem) of the innovation processes.

(2) *Criterion 2:* they use everything-mining techniques in the analysis of the innovation processes.

(3) *Criterion 3:* they study the definition of the innovation problem from the customer's or organization's perspectives.

(4) *Criterion 4:* they consider different aspects of the problem (impact, where occurs, etc.)

In Table 7, a qualitative comparison with related works is made, based on previous criteria.

As shown in Table 7, current papers did not satisfy all the criteria. Specifically, in criterion 1, our research is the only one that automates the innovation processes, in this case, using the ACODAT concept. For this automation, paradigms such as multi-agent systems can be used in conjunction with our ACODAT architecture to model the entire innovation process [24].

For criterion 2, Ossi *et al.* [19], Qin *et al.* [25], Garcia *et al.* [35] worked on the innovation based on data mining. The basis of our proposal is autonomous decisions based on knowledge models from the data extracted from market studies, internal databases, social networks, etc. Thus, this work is based on everything mining techniques. Similarly [13, 14] present autonomic cycles for self-configuration, self-optimization and self-healing during the manufacturing process based on everything mining techniques.

For criterion 3, Kritsadee *et al.* [21] tested a model of factors affecting the innovativeness of SMEs. They analyze products, processes, as well as organizational and marketing innovation. Stoettrup *et al.* [22] investigated those parameters in innovation processes and, in particular, their influence on innovation outcomes in the context of smart manufacturing. Our paper is the only one that proposes the automation of the innovation problem definition using autonomic cycles.

Finally, for criterion 4, our proposal is the only one that evaluates different aspects of an innovation problem, such as its impact on an MSME, among other aspects.

9. Conclusion

This paper proposes the automation of the innovation process in MSMEs, through the definition of ACODATs. Also, the paper applies one of the ACODAT (for the definition of the innovation problem) in an MSME, in the "RAMARA jeans" store. Our ACODATs use different data sources to build knowledge models about the innovation process (e.g. predictive and descriptive models). Through the use of our ACODATs in the innovation process, it is possible to generate knowledge for the organization, not only to identify a problem, but also, to identify where it happened, when it happened and the impact it has on the organization. Particularly, the ACODAT for the definition of the innovation problem is essential in an innovation process because it allows the organization to identify opportunities for improvement.

On the other hand, there are many data sources that companies have but do not know how to use and get the most out of them. Specifically, the multidimensional data model defined for the ACODATs determines the required information from the organization and the context. With this information, it is possible to analyze it in real time to support the decision-making process based on data, and generate useful information for the organization.

Work	Criterion 1	Criterion 2	Criterion 3	Criterion 4
[13]	✗	✓	✗	✗
[14]	✗	✓	✗	✗
[19]	✗	✓	✗	✗
[21]	✗	✗	✓	✗
[22]	✗	✗	✓	✗
[25]	✗	✓	✗	✗
[35]	✗	✓	✗	✗
This work	✓	✓	✓	✓

Table 7.
Comparison with previous works

The results of the case study allow concluding that it is feasible to use our ACODATs to automate the model of the innovation process proposed in section 3.3. The preliminary results show its utility to identify the problems that potentially will be sources of innovation processes in the organization. These preliminary results encourage the continuation of the application of the rest of ACODATs, in order to automate the innovation processes in the organizations using artificial intelligence techniques.

For future works, it is necessary to implement all ACODATs in a real scenario to verify their functionalities. To do this, it is necessary to do a detailed design of the rest of ACODATs. Also, it is important to determine the most important knowledge models required in the ACODAT for the definition of the innovation problem. Once determined, it is important to define the relevant everything mining techniques required for their implementations, such as data and process mining tasks.

References

1. Lim W, Gupta S, Aggarwal A, Paul J, Sadhna P. How do digital natives perceive and react toward online advertising? Implications for SMEs. J Strateg Marketing. 2021: 1-35.
2. Rao P, Kumar S, Chavan M, Lim W. A systematic literature review on SME financing: trends and future directions. J Small Business Management. 2021: 1-31.
3. Lim W. The quarantine economy: the case of COVID-19 and Malaysia. In: Lim W, Cheong H, Kaur S, editors. COVID-19, business, and economy in Malaysia. New York: Routledge; 2022. p. 3-23.
4. Mello S, Tomei P. The impact of the COVID-19 pandemic on expatriates: a pathway to work-life harmony?. Glob Business Organizational Excell. 2021; 40(5): 6-22.
5. Bretas V, Alon I. The impact of COVID-19 on franchising in emerging markets: an example from Brazil. Glob Business Organizational Excell. 2020; 39(6): 6-16.
6. Lim W. History, lessons, and ways forward from the COVID-19 pandemic. Int J Qual Innovation. 2021; 5(2): 101-8.
7. Castela B, Ferreira F, Ferreira J, Marques C. Assessing the innovation capability of small- and medium-sized enterprises using a non-parametric and integrative approach. Management Decis. 2018; 56(6): 1365-83.
8. Bagheri M, Mitchelmore S, Bamiatzi V, Nikolopoulos K. Internationalization orientation in SMEs: the mediating role of technological innovation. J Int Management. 2019; 25(1): 121-39.
9. Ortiz M, Joyanes L, Giraldo L. The marketing challenges in the big data age. E-Ciencias de la Información. 2015; 6(1): 16-45. 2015.
10. Sánchez M, Aguilar J, Cordero J, Valdiviezo P. Basic features of a reflective middleware for intelligent learning environment in the cloud (IECL). proceeding Asia-Pacific Conference on Computer Aided System Engineering. (APCASE); 2015.
11. Sánchez M, Aguilar J, Cordero J, Valdiviezo-Díaz P, Barba-Guamán L, Chamba-Eras L. Cloud computing in smart educational environments: application in learning analytics as service. In: Rocha Á, Correia A, Adeli H, Reis L, Mendonça Teixeira M, editors. New advances in information systems and technologies. Advances in intelligent systems and computing, 444; 2016. p. 993-1002.
12. Aguilar J, Buendia O, Pinto A, Gutiérrez J. Social learning analytics for determining learning styles in a smart classroom. Interactive Learn Environments. 2022; 30(2): 245-61.
13. Aguilar J, Garces-Jimenez A, R-Moreno MD, García R. A systematic literature review on the use of artificial intelligence in energy self-management in smart buildings. Renew Sustainable Energy Rev. 2021; 151.
14. Lopez C, Aguilar J, Santorum M. Autonomous VOs management based on industry 4.0: a systematic literature review. J Intell Manuf; 147: 2021.

15. Sánchez M, Exposito E, Aguilar J. Implementing self-* autonomic properties in self-coordinated manufacturing processes for the Industry 4.0 context. Comput Industry. 2020; 121.

16. Pacheco F, Rangel C, Aguilar J, Cerrada M, Altamiranda J. Methodological framework for data processing based on the Data Science paradigm. In: proceeding XL Latin American Computing Conference. (CLEI); 2014.

17. Puerto E, Aguilar J, López C, Chávez D. Using multilayer fuzzy cognitive maps to diagnose autism spectrum disorder. Appl Soft Comput. 2019; 75: 58-71.

18. Aguilar J. A general ant colony model to solve combinatorial optimization problems. Revista Colombiana De Computación. 2001; 2(1): 7-18.

19. Ossi Y, Jukka S, Porras J, Vesa H. Innovation capabilities as a mediator between big data and business model. J Enterprise Transformation. 2019; 1: 1-18.

20. Goodell J, Kumar S, Lim W, Pattnaik D. Artificial intelligence and machine learning in finance: identifying foundations, themes, and research clusters from bibliometric analysis. J Behav Exp Finance. 2021; 32: 100577.

21. Kritsadee P, Sanguan L, Somnuk A. Factor affecting innovativeness of small and medium enterprises in the five southern border provinces. Kasetsart J Social Sci. 2017; 38(3): 204-11.

22. Stoettrup M, Heidemann A. Design parameters for smart manufacturing innovation processes. Proced Coll Int pour la Recherche en Productique. 2020; 93: 365-70.

23. Gutiérrez A, Aguilar J, Montoya E, Ortega A. Intelligent systems used in the Micro, medium and small enterprises to improve innovation capabilities in the textile industry – a systematic literature review. Int J Entrepreneurship; 25(5S): 2021.

24. Aguilar J, Bessembel I, Cerrada M, Hidrobo F, Narciso F. Una Metodología para el Modelado de Sistemas de Ingeniería Orientado a Agentes Inteligencia Artificial. Revista Iberoamericana de Inteligencia Artif. 2008; 12(38): 39-60.

25. Qin J, Liu Y, Grosvenor R. A categorical framework of manufacturing for industry 4.0 and beyond. Proced Coll Int pour la Recherche en Productique. 2016; 52: 173-8.

26. Chopra M, Saini N, Kumar S, Varma A, Mangla S, Lim W. Past, present, and future of knowledge management for business sustainability. J Clean Prod. 2021; 328: 129592.

27. Lim W, Ciasullo M, Douglas A, Kumar S. Environmental social governance (ESG) and total quality management (TQM): a multi-study meta-systematic review. Total Qual Management Business Excell. 2022; 1-23.

28. Kephart J, Chess D. The vision of autonomic computing. Computer. 2003; 36(1): 41-50.

29. Riofrio G, Encalada E, Guamán D, Aguilar J. Business intelligence applied to learning analytics in student-centered learning processes. In: proceeding 2015 Latin American Computing Conference. (CLEI); 2015.

30. Morales L, Ouedraogo C, Aguilar J, Chassot C, Medjiah S, Drira K. Experimental comparison of the diagnostic capabilities of classification and clustering algorithms for the QoS management in an autonomic IoT platform. Serv Oriented Comput Appl. 2019; 13(3): 199-219.

31. Aguilar J, Garcès-Jimènez A, Gallego-Salvador N, De Mesa J. G, Gomez-Pulido J, García-Tejedor J. Autonomic management architecture for multi-HVAC systems in smart buildings. IEEE Access. 2019; 7: 123402-15.

32. Vizcarrondo J, Aguilar J, Exposito E, Subias A. MAPE-K as a service-oriented architecture. IEEE Latinoamerica Trans. 2017; 15(6): 1163-75.

33. Fornieles Sánchez R. De Lasswell a Gorgias: los orígenes de un paradigma. Estudios sobre el Mensaje Periodístico. 2012; 18(2): 739-55.

34. Gaind B, Varshney N, Goel S, Mondal A. Identifying short-term interests from mobile app adoption pattern. Computación y Sistemas. 2019; 23(3): 829-39.

35. Garcia J, Delsing J. Autonomous production workstation operation, reconfiguration and synchronization. Proced Manufacturing. 2019; 39: 226-34.

Supplementary material

Supplementary material is available at: https://github.com/gistag/Supplementary-Material/tree/main

Corresponding author

Jose Aguilar can be contacted at: aguilar@ula.ve

Appendix C

Sentiment analysis in the social networks for the definition of innovation problems in organizations

Sentiment analysis in the social networks for the definition of innovation problems in organizations

Ana Gutiérrez[1,2,5], Jose Aguilar[2,3,4,5], Ana Ortega[2], Edwin Montoya[2,5]

[1] GIA, Universidad Francisco de Paula Santander, Av. Gran Colombia, Cúcuta, Colombia

[2] GIDITICS, Universidad EAFIT, Medellín, Colombia

[3] CEMISID, Universidad de Los Andes, Núcleo La Hechicera, Mérida, Venezuela

[4] Universidad de Alcalá, Dpto. de Automática, Alcalá de Henares, Spain

[5] Tepuy R+D Group. Artificial Intelligence Software Development. Mérida, Venezuela

Abstract

In recent years, on social networks, people can express their thoughts, write their emotions, and generate content about many topics. In order to discover the value of this information, in this article, we propose a sentiment analysis system for the social networks based on the "autonomic cycle of data analysis tasks" paradigm, to define innovation problems in an organization. This autonomic cycle has a set of tasks used for gathering and handling large unstructured data from social media, for the definition of innovation problems through sentiment analysis. The tasks of the autonomic cycle filter the negative tweets and determine their keywords. Then, they calculate the groups/clusters in the negative tweets, and for each cluster, they analyze each centroid to obtain additional information to answer the 5 questions of the W model (What, Where, When, Why, Who) that describe a problem. Particularly, considering this information on the clusters, the last task defines customer-oriented innovation problems. Finally, this article analyzes a case about tweets in a fashion enterprise, with results very encouraging.

Keywords: Innovation Problems, Sentiment Analysis, Social networks, Autonomic Computing

1. Introduction

Sentiment analysis on social networks allows evaluating the opinions disseminated through this media. The huge volume of information (texts, images, etc.) in the social networks requires tools to automatically process these messages without losing reliability. In the context of marketing, the new trends use social networks to communicate corporate actions, mainly due to their great potential to establish relationships with customers through them [1]. Social networks facilitate the communication efforts of firms to transmit their values, attract new consumers, get feedback from discontent customers, or provide selling info [2].

Many studies have described the importance of innovation in the sustainability of an organization [3]. In general, innovation is considered vital for the survival and success of companies because of the need to evolve rapidly to follow the requirements of the environment [4]. Innovation plays a critical role in the world economy [5], it is associated with the persistent growth in companies, and can create new opportunities [6]. Fundamentally, innovation is a procedure of producing and mixing thoughts in order to solve a future problem [5]. Innovation is likewise related to issue solving, and may be the answer to a new issue or a new resolution technique to an elder issue [7].

But, before solving a problem, we must define the problem. The definition of the problem determines the scope of an innovation process. Therefore, an adequate and comprehensive definition of the problem is a key phase to be considered in any kind of innovation process [8, 9]. Thus, the analysis of the innovation process within a company with special emphasis on the phases of searching for ideas for the launch and marketing of a new product, is what is behind the definition of an innovation problem. It is during these phases that it is possible to have consumer participation through social networks contributing ideas, opinions, and supporting the launch and marketing. Social networks are becoming platforms of communication with consumers and, moreover, for supporting company innovation [10].

According to the above, it is possible to affirm that social networks are useful for the company to have extra sources of information, which will allow identifying and defining problems of innovation. In this paper, we present an autonomous cycle of data analysis tasks (ACODAT) for the definition of innovation problems based on sentiment analysis in social networks. We identify the innovation problems using the ACODAT concept [10][11][12], and particularly, a previous generic ACODAT whose objective is the definition of an innovation problem [13]. In this paper, this autonomous cycle is instantiated to analyze negative tweets, such that each data analysis task is defined to answer the questions of the W model (What, Who, Where, When and Why) [29] with the aim of defining the innovation problem from the negative customer comments made on Twitter. The major contributions proposed in this paper are:

1. Carry out one instance of ACODAT in order to define one innovation problem using sentiment analysis.
2. Adapt the tasks of this ACODAT with the IA techniques required for the innovation problem definition using the social networks as information sources.
3. Describe a case study of this ACODAT in the textile industry of fashion using Twitter as the social network.

The structure of this work is as follows. Section II describes the associated works. In Section III, the theoretical framework of this work is described, specifically, the ACODAT for the definition of innovation problems. Section IV describes the instantiation of this ACODAT using social networks, in particular Twitter. Section V describes the textile case study, including the experiments and analysis of results. Section VI purposes criteria to compare our proposal with other works. Finally, the main conclusions and next works are exposed.

4. Related works

In this section, we describe the primary latest papers associated with our approach. These works carry out a sentiment analysis in the social networks and/or study innovation problems. First, we present the papers about sentiment analysis for innovation contexts [1][14][15][20]. Second, we present the papers about the definition of the innovation problems in the context of innovation processes [21][22][23][24][25].

The paper of Hutto and Gilbert [14] collects and manages massive unstructured information from social media for sentiment evaluation for figuring out services. Particularly, they define a technique described by a set of rules built from the sentiment assessment of customers by service classes, grouped by domains. Saura and Palos [1] analyzed the sentiments (positive, negative, and neutral) of the offers published on Twitter through the hashtag #BlackFriday. They identified the sentiment of the offers published on the social network to promote the #BlackFriday event. Using this information, they identify strategies to improve the commercialization and marketing of companies [26].

In the article [15], Kumar and Jain define an automatic assessment approach considering the sentiment analysis of student opinions. The comments from the students are collected and the system automatically identifies their sentiments, as well as relevant features of a product or service employed. Khramoin et al. [20] describe a sentiment evaluation that uses a set of rules for the assessment of the public's mindset toward innovations. The sentiment analysis considers textual content fragments, called emotive constructions, during the analysis. This paper confirmed that sentiment analysis can improve the manual analysis of professionals. Zhou et al. [22] define a two-layer model to describe the needs of customers. The first layer is based on sentiment analysis to determine customers. The second layer detects the characteristics of these customer needs, by analyzing semantic differences and similarities.

In the work (16], the authors examine the impact of media on public opinion. The article describes how the media impact public opinion. They indicate that this control and influence considers cognitive, psychological, physiological, and technological elements. They conclude that the media shape social constructs and behaviors, with fundamental consequences for the future of today's society. Based on these results, Sanchez et al. [17] define a model that describes the conformation of public opinion from the angle of opinion mining, through the use of the Multilevel Fuzzy Cognitive Map technique [18][19]. They study the abilities of the approach for both, knowledge extraction and knowledge description for opinions on different topics.

In general, Artificial Intelligence (AI) has the potential to improve the innovation process. For example, AI can become aware of patterns, synthesize information, draw conclusions, make predictions, or perform problem-solving tasks [23]. Haefner et al. [21] review the use of AI technologies for innovation management. They outline a framework to define where AI, and particularly, their use for the digital transformation of innovation. Kakatkar et al. [24] describe strategies for the generation of information and models for innovation analysis using data-driven approaches. They show two case studies, certainly considered one among a German non-public care product producer that used AI algorithms to examine chats and extract purchaser needs. Another case is an American semiconductor chip manufacturer that used AI to identify issues related to their products. In the paper [25], Tsung-Ting et al. integrate social network analysis (SNA) and natural language processing (NLP) to estimate the propagation of a new subject matter and define a framework for this problem. They use latent semantic information between topics, users, and social connections for the predictions.

According to the previous works, there are articles related to sentiment analysis applied in different innovative contexts such as improving product promotion strategies [1], evaluating public opinions [17] [20], and identifying customer needs [22], among others. However, there are no works defining innovation problems using sentiment analysis. Additionally, the ACODAT concept has been used in other domains [11][12] [27][28], and a preliminary work has built a generic ACODAT to autonomously define innovation issues [13]. This paper is based on this ACODAT and suits it for the sentiment analysis in social networks. These are the principle variations in our technique concerning preceding works.

1. Autonomic Cycle to analyze innovation problems

In this work, we consider one of the ACODAT described in [13]. The authors of [13] propose ACODAT to enhance the innovation process. They design 3 ACODATs: one to specify the needs, another to define innovation problems, and finally, one to evaluate the results. The second one is studied in this work, which will be instanced in the context of sentiment analysis.

The relevance of an ACODAT to define Innovation problems (ACIP001) is because it allows its characterization [13]. In general, this ACODAT has a set of data analysis tasks, which use mining techniques to acquire useful data to build the innovation problem statement. We use the 5W approach to build this cycle (see Figure 1) [29].

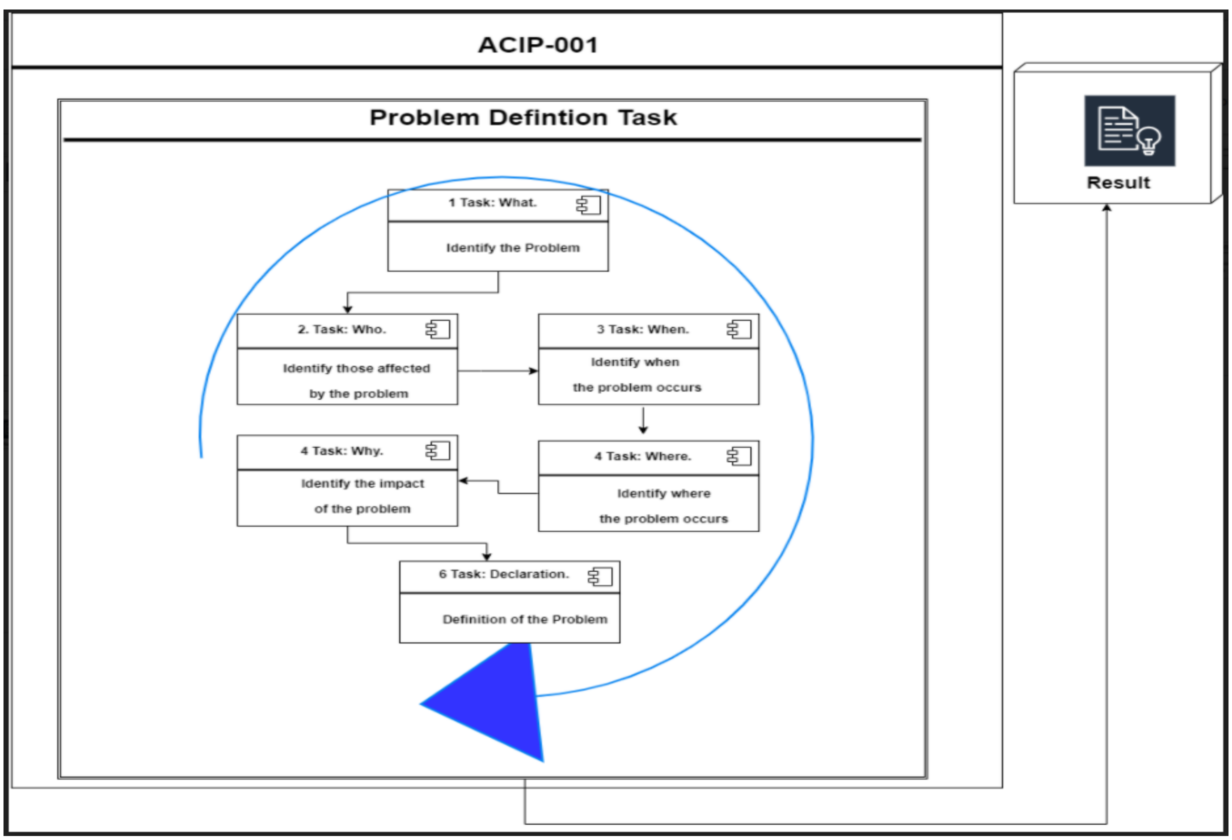

Figure 1. Structure of the ACIP-001 [13].

Table 1 suggests the overall description of every task of this ACODAT described in [13], based totally on the 5W model:

Task Name	Knowledge Models
1. What: it identifies the potential innovation problem	Descriptive and Detection Models
2. Who: it establishes those affected by the potential innovation problem	Descriptive Model
3. When: it determines when the potential innovation problem occurs	Detection and Predictive Models
4. Where: it identifies where the potential innovation problem occurs	Diagnostic and Predictive Models
5. Why: it determines the impact of the potential innovation problem	Diagnostic and Predictive Models
6. Declaration: it defines the innovation problem	NLP

Table 1. Description of the Tasks of ACIP-001

Now, we describe each task.

1. What: in this task is identified the trouble with the data sources. An example of a problem is the purchaser lawsuits. The goal of this task is to decide the prevalence of innovation troubles (i.e., it's far vital to create an authentic solution). This task uses descriptive and detection models to become aware of the innovation problem.

2. Who: This task identifies who is affected by the previous innovation problem (e.g. customers, organizations, or specific groups). Due to that, it uses descriptive models.

3. When: this task establishes when the problem will occur, for which it uses detection or prediction models.

4. Where In this task is defined where the problem occurs, for which it uses data sources about geo-localization and diagnosis models.

5. Why: In this task is determined the relevance of the problem.

6. Declaration: this task defines the problem. This task is able to use NLP strategies to define the statement problem.

2. Instantiation of ACIP-001 based on Social Networks

In this section, we describe the Instantiation of ACIP-001 using Social Networks. Based on the ACODAT for the "Definition of the innovation problem" defined in [13], its instantiation is defined by the macro-algorithm in Figure 2.

Input: Dataset

Output: Definition of Problem

1. Apply Sentiment Analysis
    1. Filter the data source
    2. Standardize information
    3. Preprocess information

Figure 2. SASN-IPO macro-algorithm

This macro-algorithm describes the different functions that must be covered by the tasks of the autonomous cycle for the SNA. Each of them is explained below:

1. Apply sentiment analysis to detect the polarities (positive, negative, neutral, or compound) in a text, paragraph, sentence, or clause. For this task, several steps are required:

    a. Standardize the information to normalize.

    b. Preprocess the information to transform the raw data for analysis and modeling algorithms. Particularly, it carries out NLP processes such as Tokenize, Remove Clean characters and Stop Words, and Stemming/Lemmatization.

    c. Obtain keywords to categorize each data source.

    d. Filter the information in the data source to use the more relevant for our context (in our case, only negative tweets).

2. Carry out a clustering analysis to respond to the answer of the 5W models. In this step are used the first tasks of ACIP-001.

3. Generate the innovation problem statement. This step is the last task of ACIP-001 and carries out a natural language generation using the results of the previous tasks for the definition of the innovation problem.

Table 2 shows a modification of ACIP-001 based on the macro-algorithm of Figure 2.

Autonomous Cycle Task	SASN-IPO macro-algorithm	Data Source
Task 0. Sentiment analysis	1.a. Standardize information 1.b. Preprocess information 1.c. Keyword acquisition 1.d. Filter the information	Twitter
Task 1. What	2.1 Clustering analysis based on Keywords	Keywords
Task 2. Who	2.2 Clustering analysis based on ids	Id users
Task 3. When	2.3 Clustering analysis based on the date	Date, month, working day
Task 4. Where	2.4 Clustering analysis based on location	Localization
Task 5. Why	2.5 Clustering analysis based on the intensity of feeling on the tweets	Compound
Task 6. Declaration	3. Definition of the problem	Compilation of previous answers

Table 2. Autonomous Cycle Task associate SASN-IPO macro-algorithm

Now, we describe each task.

Task 0. Sentiment analysis. This task is defined by the three next sub-algorithms to obtain the data sources for the next tasks. This task carries out an SNA process, which in our case is linked to Twitter.

Standardize information: this step normalizes the information to be used. Figure 3 shows the sub-algorithm that defines the steps to extract the information from the Twitter dataset. For example, the year is extracted from the date value in the dataset. From the value of the date also is extracted the month and hour is extracted. If the hour is less than noon (12) the day column is created with a value of 1 (mean morning day), else it will be 2 (mean afternoon day). Similarly, it is done with the rest of the information of interest contained in the tweets to answer the questions of the 5W model (users, with a value of 1 for the internal users, and otherwise 2 for the external users; and location). Now, a very important information to extract is the text of the tweets, since it will be used to carry out the sentiment analysis on it.

Input: Dataset

Output: Standardize Dataset

1. Standardize Information

   a. Extract from the date field the value of the year

   b. Extract from the date field the value of the month

   c. Extract from the date field the value of the hour

   d. If hour < 12:00pm : a. Set working day = 1

      else b. set working day= 2

Figure 3. Standardize information sub-algorithm

Preprocess information: in this step is carried out the transformation of the data from the text of the tweets for the next analysis. Particularly, it tokenizes, removes Stop Words, cleans special characters, and stems/lemmatizes the text. Figure 4 presents the sub-algorithm to preprocess tweets (a typical NLP procedure).

Input: Text Twitter

Output: Clean Text Twitter

1. Input Text Twitter

2. Remove Stop Words

3. Remove URLs

4. Remove special characters

Figure 4. Preprocess information sub-algorithm

Keyword acquisition: In this step, the Twitter texts are preprocessed to obtain the keywords. With the clean texts of each Twitter are determined its keywords using the tf-idf metric. This metric increases according to the frequency of a word in the content of a Twitter, and decreases by the frequency of the word in the group of Tweets. Thus, the words with the highest value will be the keywords of each text of the analyzed Tweets.

Filter the information: This is the sentiment analysis phase. According to the keywords previously identified by text, the polarities of each one are determined, by their intensities. From there, an average per text is calculated (it is the average of the polarities of its keywords). To determine the polarity of each word in the tweets, the vaderSentiment library is used [14]. This library defines the polarity of a word, classifying it into positive, negative, neutral, or compound (hybrid emotion). Additionally, this library defines the compound-score value of each text, which is the average polarity of all words in the text. In Figure 5 is shown the macro-algorithm to carry out the sentiment analysis of the dataset. If the compound-score value is less than 0 then Twitter is negative, otherwise, Twitter is positive. The next step is to filter the negative tweets, this new subset of tweets will be the source of information to continue with the next tasks.

Input: Initial Dataset

Output: Negative Dataset

1. Apply sentiment analysis

2. If Compound_Score < 0

    a. set Twitter Negative

Figure 5. Sentiment Analysis sub-algorithm

Thus, the result of this task is all the negative Tweets, which will be the data source for task 1.

Task 1. What: Identify the problem: For this task is necessary to build groups where centroids serve to define innovation problems. Figure 6 presents the sub-algorithm that defines the clustering problem based on the keywords of the tweets. Particularly, step 1 calls the K-means clustering algorithm.

Input: Tweets defined by their keywords

Output: Clusters

1. Call the Clustering technique

Figure 6. Keyword-driven Clustering sub-algorithm

The silhouette metric has been used with the K-means clustering algorithm to determine the number k [30]. The silhouette value becomes higher as two data within the same cluster are closer. It also becomes higher as two data in different clusters are farther. Typically, a silhouette value ranges from − 1 to 1, where a high value indicates that data are well matched to their own cluster and poorly matched to neighboring clusters. Generally, the silhouette value is greater than 0.5 means that clustering results are good [30].

Finally, this task analyzes the keywords that describe the centroids of each cluster, to determine the answer to the "What" of the 5W model. In summary Figure 7 shows the main steps to determine the "What" using the centroids.

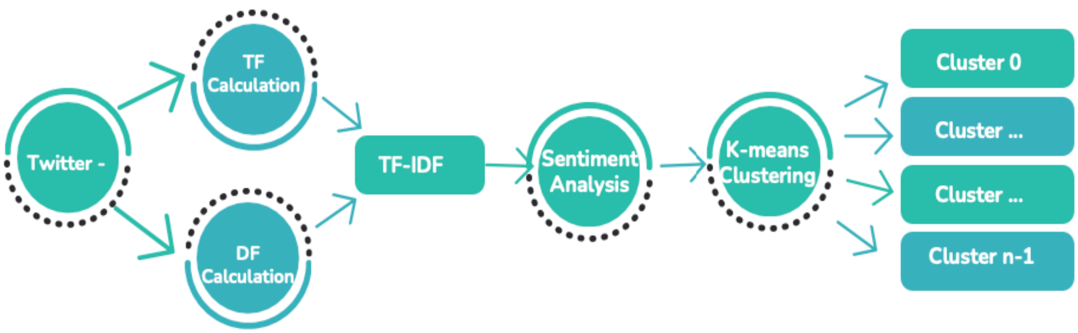

Figure 7. Determination of the "What" of the 5W model

Task 2. Who: Identify the ones affected by the problem: This task uses the information of the previous task to become aware of who the problem affects. For this, for each cluster obtained in task 1, a new clustering is carried out based on who the Twitter users are (see Figure 8). In this case, the persons could be an employee of the company, physical consumers, or online consumers.

Input: Twitter Clusters and user-Ids

Output: Who

1. Call the Clustering technique based on user information

Figure 8. User-driven Clustering sub-algorithm

Task 3. When: Identify when the problem occurs: This task uses the information collected in task 1 to identify when the problem occurs. In this case, a new grouping is obtained for each cluster of the first task, using the date as similarity criteria (see Figure 9).

Input: Twitter Clusters and date

Output: When

1. Call the Clustering technique based on date information

Figure 9. Data-driven Clustering sub-algorithm

Task 4. Where: Identify where the problem occurs: This task uses the information collected in task 1 to identify where the problem occurs. In this case, again the clusters of the first task are re-clustered using the geo-localization of the customers according to the social network (see Figure 10).

Input: Twitter Clusters and Geo-localization

Output: Where

1. Call the Clustering technique based on geo-localization information of the customers

Figure 10. Geo-localization-driven Clustering sub-algorithm

Task 5. Why: Identify the impact of the problem: This task uses the information collected in task 0 to identify the impact of the problem. In this case, we use the compound score value of the task of sentiment analysis to determine the impact of the problem (see Figure 11).

Input: Twitter Clusters and compound_score

Output: Why

1. Get compound_score value

if compound_score is between [-0,01 , -0,33):

a. Set why = High

else if compound_score is between [-0,33 , -0,66):

Figure 11. Impact Identification sub-algorithm

Task 6. Declaration: Definition of the problem: This task uses all the results of the previous tasks to define the innovation problem. We use the keywords and the next pattern to build the sentences that defined the innovation problem:

<Who(user)> + "published" + <When(date)> + "from" + <Where(localization)> + "about" + <What(keywords)> + "such that the impact is" + <why>

That corresponds with each one of the results of the previous tasks.

3. Experiment and Analysis of results

This section presents the experimental context for the instantiation of the Autonomous Cycle for the Definition of the Innovation Problem (ACIP001 Definition of the Problem) using social networks.

1. Experimental Context.

In this case study, we used a dataset from the "OFFCORSS" company in Colombia. This company is dedicated to the manufacture, sale and marketing of all kinds of clothes for babies, girls and boys such as T-shirts, shirts, skirts, shorts, jackets, dresses, jeans, pants, swimwear, shoes, underwear, pajamas, among other products. Its goal is to offer a first-class and pleasant product. The company currently has social networks on Facebook, like @offcorss; Instagram, like @offcorsscolombia; Pinterest, like @offcorss; Twitter, like @OFFCORSS, and Youtube. It also has a team dedicated to virtual sales of products nationwide to attend to all requests, doubts, and questions of its customers.

For this case, we use its social network on Twitter (@OFFCORSS). The data of the tweets are the message (text), with the information about the user who sends it (id_user, screen_name), the date of creation (created at), url_Tweet, etc. The records were provided by the OFFCORSS Company and correspond to tweets from the period of June-August 2021.

2. ACIP-001 in the case study

In this section, we describe each of the tasks of the cycle autonomous for the definition of the innovation problem.

Task 0. Sentiment analysis. This task gets sources of data, normalizes them, executes an NLP process, and finally, carries out the sentiment analysis. The company provided a total of 695 tweets. The dataset was made up of raw messages, so it needed prior cleaning before labeling it. Table 3 shows the *preprocessing step*, where the irrelevant information was removed from the tweets. Some of this information is data like URLs, RT, mentions (@), hashtags(#), punctuation, and accents.

	TweetE	compound_Score	...	Sent_comp	cluster	TweetE2	text_clean	text_clean2	word_count	avg_word	stopwords	hashtags	numerics
	@DanielSamperO @fdbedout there is no gentleman...	-0.2960	...	2	2	@DanielSamperO @fdbedout gentleman size that's...	@danielsampero @fdbedout gentleman size that's...	gentleman size that's small go offcorss	6	5.666667	0	0	0
	RT @Dantestereo Putting the former prisoner in...	-0.9231	...	2	2	RT @Dantestereo Putting former prisoner prison...	rt @dantestereo putting former prisoner prison...	rt putting former prisoner prison bribing witn...	13	7.000000	0	0	0
	Putting the ex-prison in prison for bribing wi...	-0.8689	...	2	2	Putting ex-prison prison bribing witnesses use...	putting ex-prison prison bribing witnesses use...	putting ex-prison prison bribing witnesses use...	11	7.636364	0	0	0
	@OFFCORSS Perhaps if today in the world and in...	-0.9074	...	2	2	@OFFCORSS Perhaps today world country afflicte...	@offcorss perhaps today world country afflicte...	perhaps today world country afflicted pandemic...	25	6.280000	0	0	0

Table 3. Results of the Preprocess information step

Now, we *determine the keywords* using the tf-idf value for each word in the text of the tweets. First, we calculate tf, idf, and tf-idf. Then, we select the keywords, which are the words with a tf-idf value superior to 3. For example, in Table 4, for a given Twitter we show the tf-idf value of the words on its text. Priority is given to words with the highest value of tf-idf (Step 1). The rest of the keywords (steps 2 and 3) will be explained in the next step.

	words	tf	idf	tfidf	negative	compound	polarity	
2	habeas	2	3.496508	6.993015	0.0	0.0000	0.000000	
1	report	2	3.496508	6.993015	0.0	0.0000	0.000000	Step1
0	data	2	3.496508	6.993015	0.0	0.0000	0.000000	
11	negative	1	2.803360	2.803360	1.0	-0.5719	-0.300000	Step 2
7	violated	1	2.803360	2.803360	1.0	-0.5267	0.000000	
17	report.	1	3.496508	3.496508	0.0	0.0000	0.000000	
3	current	1	3.496508	3.496508	0.0	0.0000	0.000000	
4	notified	1	3.496508	3.496508	0.0	0.0000	0.000000	
5	withdrawn	1	3.496508	3.496508	0.0	0.0000	0.000000	
9	otherwise	1	3.496508	3.496508	0.0	0.0000	0.000000	Step 3
10	receive	1	3.496508	3.496508	0.0	0.0000	0.000000	
12	person	1	3.496508	3.496508	0.0	0.0000	0.000000	
14	before	1	3.496508	3.496508	0.0	0.0000	0.000000	
15	requires	1	3.496508	3.496508	0.0	0.0000	0.000000	
16	law	1	3.496508	3.496508	0.0	0.0000	0.000000	
6	must	1	2.803360	2.803360	0.0	0.0000	0.000000	
8	right	1	2.397895	2.397895	0.0	0.0000	0.285714	
13	making	1	2.397895	2.397895	0.0	0.0000	0.000000	

Table 4. Selection of keywords for each tweet.

For the *sentiment analysis* was used the vaderSentiment library in each Twitter register (the text column) to determine the sentiment of each one. This library computes the compound_score of each text. Table 5 presents a partial view of the results of the sentiment analysis for this dataset. Remember that the tweets are negative when the value of compound_score is < 0. A total of 233 negative tweets were obtained.

Usuario	Fecha Creacion	Tweet	Sentimiento	UrlenTweet	UrlTweet	TweetE	compound_Score	...
Javier Neira	Aug 5 2020 04:53 PM	@DanielSamperO @fdbedout de esa talla no hay c...	negative	NaN	https://twitter.com/Javier_Neira_/status/12910...	@DanielSamperO @fdbedout there is no gentleman...	-0.2960	...
Scofield	Aug 3 2020 11:44 PM	RT @Dantestereo: Que metan preso al ex presi p...	negative	NaN	https://twitter.com/Ivanzuluaga95/status/12904...	RT @Dantestereo Putting the former prisoner in...	-0.9231	...
Al Efe Quatro	Aug 3 2020 10:34 PM	Que metan preso al ex presi por soborno de tes...	neutral	NaN	https://twitter.com/Dantestereo/status/1290415...	Putting the ex-prison in prison for bribing wi...	-0.8689	...
Nina	Jul 30 2020 12:29 AM	@OFFCORSS Tal vez, si hoy en el mundo y en el ...	negative	NaN	https://twitter.com/Caromaticas/status/1288632...	@OFFCORSS Perhaps if today in the world and in...	-0.9074	...

Table 5. Results of the sentiment analysis

Now, we complete the list of keywords (see Table 4). For that, those words that have a negative polarity, or a compound less than zero (Step 2) are included if they have a high value in tf-idf. Finally, if the number of words selected in steps 1 and 2 is a lower number to half of all the words in the text, then in step 3 are selected the words that follow in descending order according to the tfi-df value until completing the half.

Thus, the first step selects the most relevant three words, the second step selects the most negative words considering the negative values (compound and polarity variables of the word). Then, according to the descending order of relevance of the words, more words are selected until completing the half. This is the list of keywords of each Twitter analyzed for the next tasks.

Task 1. What: Identify the problem: this task defines the clusters such that their centroids serve to define innovation problems. Thus, using the keywords in the tweets it is invoked the K-means clustering algorithm. According to the value of the silhouette metric, 8 clusters were found (see Table 6).

N clusters	**Average silhouette**
0	0.78
1	0.72
2	0.69
3	0.71
4	0.83

5	0.87
6	0.88
7	0.87

Table 6. Clusters based on the keywords in each twitter

The centroids of the clusters defined by this task can be observed in Table 7. We present the keywords of each cluster centroid with which the innovation problem should be defined.

Cluster centroids			
Cluster 0:	**Cluster 1:**	**Cluster 2:**	**Cluster 3:**
hotline	negative	worst	terrible
stolen	violated	answer	service
minutes	makes	happened	bought
waiting	responding	email	bad
follow	rating	assistance	refund
resources	platform	robbery	requested
gentleman	timely	say	claim
size	habeas	vigilance	WhatsApp
small	report	bonus	requesting
price	data	consumer	process
Cluster 4:	**Cluster 5:**	**Cluster 6:**	**Cluster 7:**
missing	meet	prison	worse
email	comes	putting	unacceptable
follow	delivered	stealing	disagreement
finally	sadly	useless	complaint
fact	quarantine	prisoner	wrong
ex	dispatched	ex	email
errors	used	errors	finally
emails	fact	finally	fact

domiciles minutes	ex errors	fact emails	ex errors

Table 7. Keywords of the centroid of each cluster

For the following tasks, grouping processes are carried out for each of the clusters obtained in this phase, using in each case different criteria of similarity (user, date, and place).

Task 2. Who: Identify the ones affected by the problem: We present the results of the groups by users in each cluster to determine the users who make the publication in the social network. Figure 12 shows a part of the cluster generated by this task. To explain this task and those that follow, we are going to analyze the information of two clusters (clusters 1 and 3). In cluster 1 there are internal users and external customers (1 represents an internal user e.g., one employee of the company and 2 corresponds to an external customer). For cluster 3, only external users published tweets.

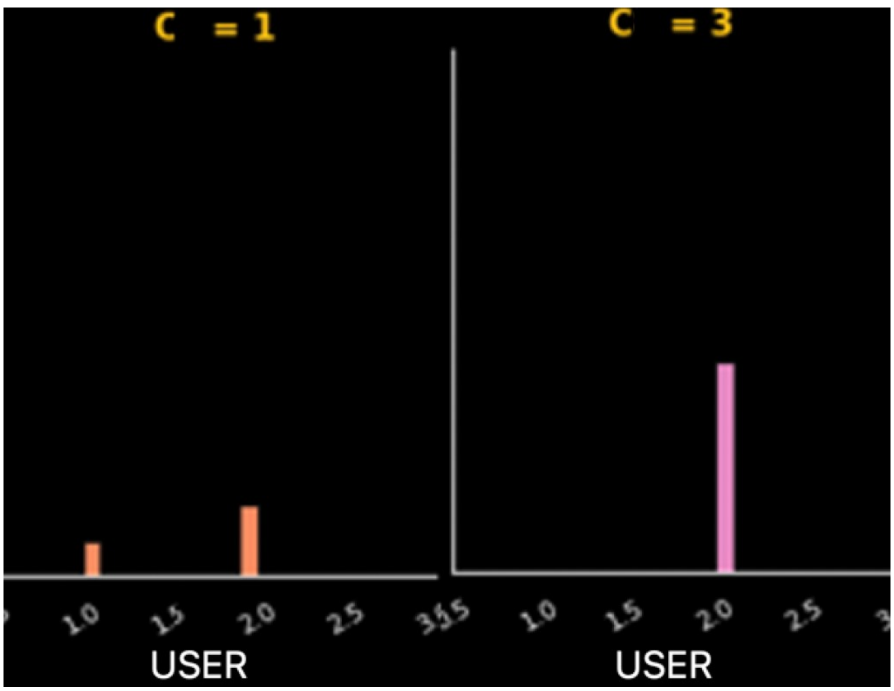

Figure 12. Partial results of Task 2

Task 3. When: Identify while the hassle occurs. Figure 13 shows a partial view of the results of task 3. It shows the months, and if the tweets were sent in the morning or afternoon. For example, for cluster 1 the date of publication of the tweets was in July (7) in the morning (1) and afternoon (2), while for cluster 3 the date of publication was in July (7) and only in the afternoon (2).

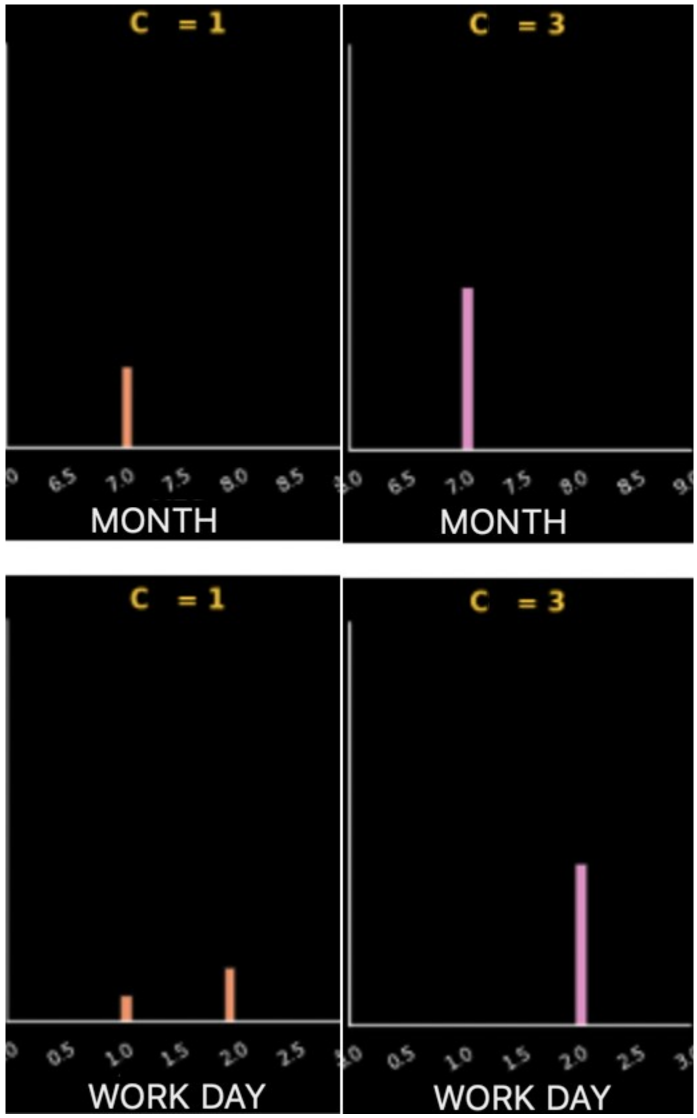

Figure 13. Partial results of Task 3.

Task 4. Where: Identify where the problem occurs. We present the result of the localization of each user into the clusters. Figure 14 shows a partial view of results based on the different geo-localizations of the users in the social network. According to the data, the localizations were in Bogota, San Andrés, Popayan, Choconta, Medellin, Montería, Piedecuesta, and some of them cannot be identified because users configured their data not to be visible. For cluster 1, the publications were made from Popayan, Medellin or the location is not known. For cluster 3, there is no record of where the publications were made because the location value is 0.

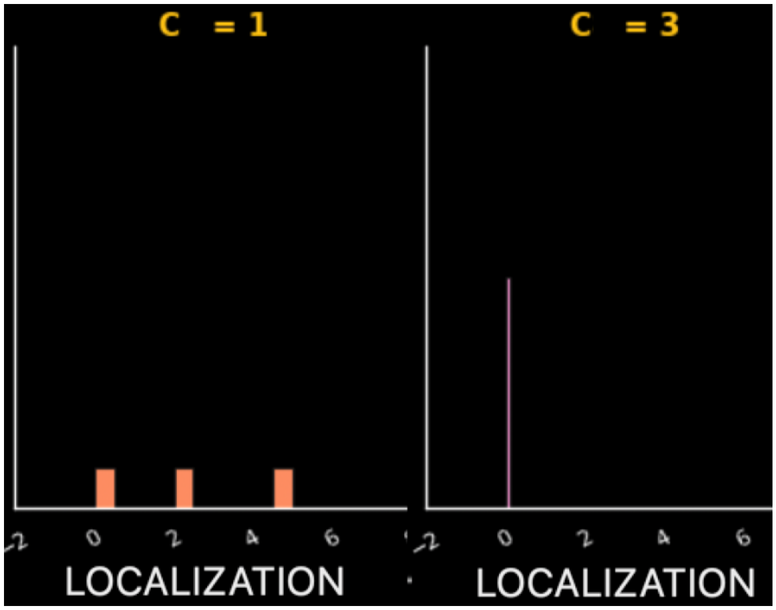

Figure 14. Partial results of Task 4

Task 5. Why: Identify the effect of the problem. This task uses the information collected in task 0 to identify the impact of the problem. It uses the result of the compound score of each user into the clusters. The value between (-0,66 to – 1] is Low impact, between (-0,33 to – 0,66] is Medium impact, and between (-0,01 to – 0,33] is High impact. In this case, Figure 14 shows for cluster 1 that the average value of the compound_score variable was -0.55, and for cluster 3 it was -0.39. Therefore, the impact for the two clusters is Medium.

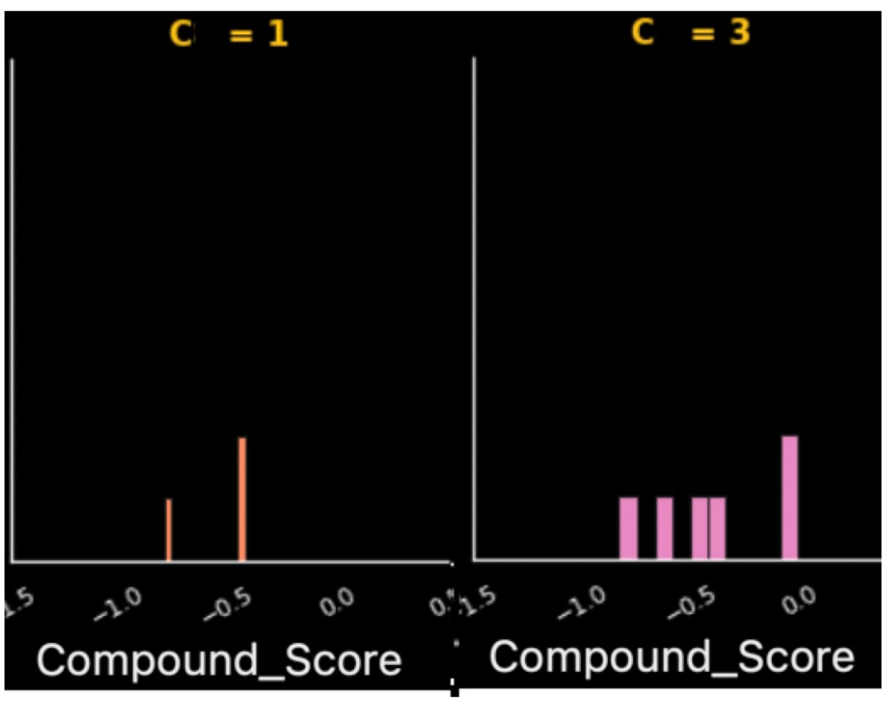

Figure 15. Partial results of Task 5

Task 6. Declaration: Definition of the problem: This task uses all the results of the previous task, to define innovation problems. We present the result for clusters 1 and 3. According to the pattern defined in section IV, Table 6 shows the statement of the innovation problem defined by cluster 1.

Task.		Results
1.	What	Negative report, responding data, etc.
2.	Who	Clothing customers and one internal user
3.	When	In the month of July 2021 in the afternoon session and one in the morning session.
4.	Where	The publications are from Popayan, Medellin and one is not possibly known.
5.	Why	The impact is Medio

Table 6. Interpretation of the information for cluster 1

The declaration of the innovation problem using Cluster 1 is:

Clothing customers and one internal user ***published*** *in July 2021 in the afternoon session and one in the morning session* ***from*** *Popayan, Medellin and one is not possibly known* ***about*** *Negative report, responding data.* ***The impact is*** *Medio.*

Maybe we can think about there is no timely response to the client. This implies that it is necessary to innovate in the ways of anticipating the demands of the clients, to give opportune answers.

Table 7 shows the statement of the innovation problem defined by cluster 3.

Task.		Results
1.	What	Terrible service, refund money, etc.
2.	Who	Clothing customers
3.	When	All published in July 2021 in the afternoon session.
4.	Where	Most users had the location disabled so it is unknown where they posted from.

5.	Why	The impact is Medio

Table 8. Interpretation of the information for cluster 3

The declaration of the innovation problem using Cluster 3 is:

*Clothing customers **published** in July 2021 in the afternoon session **from** the location is not possibly known **about** Terrible service, refund money. **The impact is** Medium.*

According to the previous sentence, maybe the shipment does not arrive, the client services never respond in a timely, and they don't refund money either. This implies that it is necessary to innovate in these services to improve customer service in general.

1. Discussion general of the results

Task 1. What? The problem arises with the "attention to online customers" of the company OFFCORSS, which provided an online sales service. The problem that requires more attention is related to the timely dispatch of the merchandise as well as the immediate response to customer complaints due to errors such as: lack of inventory, violation of consumer rights, and non-compliance with the delivery of merchandise, misleading offers, defenselessness of the client and duplicate collection, among others

Task 2. Who? The issue directly affects the client as the primary user and their family group as the secondary user. It also directly affects the company, which will be affected in its future image by being involved in the errors mentioned above.

Task 3. When? The problem arose in the year 2021, in July and August, during the pandemic.

Task 4. Where? The problem is occurring in different cities in Colombia.

Task 5. Why? It is important to solve the problem to restore the customer's credibility in the company and its online services, and to improve the company´s corporate image and brand's public perception.

Task 6. Statement of the problem: addressing an Innovative Project for virtual marketing entails a series of elements that come together from a diagnosis of the needs of the company to the solution of the problem by formulating it, and the consideration of the organizational, human, technical, and economic, among others, to solve it.

Faced with the first question: What? it can be said that by experiencing the practice of online shopping and in the face of the growing demand for online offers, a reflection is made on the skills that companies that they venture as virtual stores, to facilitate their customers access to their products or services, which must be done with ethic and responsibility. Likewise, a second reflection is worth: what for? In this case, it is necessary to solve a logistical, technological, and service provision problem, in terms of tools and resources to use, and maintain a competent staff in the dispatch, transport and collection of merchandise.

There is a third question about why do it? It arises from the need for timely attention to online customers of the OFFCORSS Company in Colombia, because there is a series of errors reported by customers through Twitter. Some of the problems are: There is no merchandise for children in some stores; there is no timely response to the client, there is a violation of consumer rights; the purchases do not arrive, and the money is not refunded; among other problems.

It is important to solve the problem to restore the credibility of the client in the company and in its online services. In the same order of ideas, facing the rest of the questions and answering the Who? The problem directly affects the client as a primary user, and their family group as secondary users. Consequently, the company will see its image damaged. When? The situation arose in 2021, in July and August, during the pandemic. Where? The diagnosis was made with the company OFFCORSS, a virtual store, located in Colombia, a situation detected through the opinions of customers through the social network Twitter.

6. Comparison with previous works

In this section, we propose several criteria to compare our proposition of sentiment analysis to analyze innovation problems using social networks with other works. We define the next criteria:

Criterion 1: They use data sources from some social networks, e.g., Twitter, Facebook, and Instagram.

Criterion 2: They develop a sentiment analysis.

Criterion 3: They analyze groups of users and interpret their information to define innovation problems.

Criterion 4: They propose autonomic processes for the self-defining of innovation problems.

Criterion 5: They are based on the 5W model to define innovation problems

In Table 10, a qualitative evaluation with related works is made, using these criteria

Work	Criterion 1	Criterion 2	Criterion 3	Criterion 4	Criterion 5
[13]	✓	✗	✗	✓	✓
[14]	✓	✓	✗	✗	✗
[22]	✓	✓	✓	✗	✗
[25]	✓	✗	✓	✗	✗
This work	✓	✓	✓	✓	✓

Table 10. Comparison with preceding works.

For criterion 1, all the works use as a source of data any social network. For criterion 2, Hutto and Gilbert [14] use sentiment analysis in microblog-like contexts. Also, Zhou, et al. [22] use sentiment analysis to identify customer needs. For criterion 3, Tsung-Ting et al. [25] estimate the propagation of a new subject and build a framework for this problem.

For criterion 4, Gutiérrez et al. [13] propose an ACODAT for this phase of the innovation process to be autonomous and self-managed in order to definite innovation problems. For criterion 5, Gutiérrez et al. [13] use the model 5W in their proposal. They purpose the sources of data, and the AI models for each one of the tasks to answer the questions of the 5W model to define innovation problems.

As shown in Table 10, current papers do not satisfy all five criteria. Specifically, in criteria 4 and 5, our research is the only one that considers, analyzes, and interprets cluster information to define problems. In this case, using an ACODAT to analyze innovation problems based on the 5W model, but from the point of view of social networks (in this paper, tweeters).

7. Conclusions

Undoubtedly, the use of digital resources and tools for the automation of the cycle of data analysis tasks is of great contribution to the identification of problems within the innovation processes of companies. Particularly, in this article, we developed a protocol for the application of the ACODAT and address a case based on real data obtained from a social network for a Colombian company dedicated to the manufacture, marketing, and sale of clothing.

In this paper each one of the tasks of the ACODAT to define innovation problems was implemented, to verify its functionality, and to evaluate the quality of the results. In this case, we have used sentiment analysis and NLP to identify where a problem happened, when it happened, and the impact it has on the organization. So the information extracted from tweets using our ACODAT allows modeling and visualizing the innovation problems.

The methodological development, which is addressed step by step in the article, provides a guide based on ACODAT theory, but also, represents an example for organizations that wish to apply the sentiment analysis process to identify problems in their innovation processes. It is also important to remark that there are many sources of data that companies don't know how to use and get the most out of, and this knowledge becomes a strategic asset that needs to be known to extract and utilize. Being that case, we have shown how social networks give useful information for the operations of an organization, but for the sake of clarity, the process developed can be implemented using a wide range of information sources. In view of the valuable results obtained, we consider that future work might explore new mechanisms that allow us to integrate additional information from the organization, for example, satisfaction surveys, requests, complaints and claims from customers, information from other social networks such as Facebook and Instagram, to be used by the autonomous cycle.

Particularly, it was observed that a Twitter customer service dataset might provide a large corpus of tweets and replies that can support innovation and can be used to improve innovation. Next works can use other social networks such as Facebook and apply the process in other industries such as airlines, and public service companies.

Finally, given the possibility of including AI in the innovation process of an organization, the question arises as to when, how, and to what extent the interactions must be between the innovation teams of the organization (for example, the innovation managers) and AI tools, to achieve maximum efficiency in said process. For instance, current analyses estimate that AI technology can improve 60% of all current occupations [31]. Consequently, we assume it critical to adopt an extra particular dialogue of AI's capacity to replace human beings inside the innovation process of organizations, especially those that have access to different sources of customer information and perception [17]

References

[1] J. Saura, A. Reyes-Menendez, and P. Palos-Sanchez, “A Sentiment Analysis on Twitter with Machine Learning: Identifying the sentiment around #BlackFriday deals,” *Journal of Open Innovation: Technology, Market, and Complexity*, vol. 3, 2019, doi: https://doi.org/10.3390/joitmc5030058.

[2] P. Paroubek and A. Pak, “Twitter as a Corpus for Sentiment Analysis and Opinion Mining,” 2010.

[3] A. Buhl *et al.*, “Design thinking for sustainability: Why and how design thinking can foster sustainability-oriented innovation development,” *Journal of Cleaner Production*, vol. 231, pp. 1248–1257, Sep. 2019, doi: 10.1016/j.jclepro.2019.05.259.

[4] L. E. Ferrer, C. Karina, P. González, I. Lilia, and M. Mendoza Vega, “Innovation as a key factor for improving the competitiveness of SMES in the department of Atlantic, Colombia,” *Dictamen Libre*, vol. 16, pp. 21–36, 2015.

[5] S. Baskaran and K. Mehta, “What is innovation anyway? Youth perspectives from resource-constrained environments,” *Technovation*, vol. 52–53, pp. 4–17, Jun. 2016, doi: 10.1016/j.technovation.2016.01.005.

[6] V. Denicolò and P. Zanchettin, “Speculative profits, innovation, and growth,” *Economic Inquiry*, vol. 55, no. 1, pp. 160–174, Jan. 2017, doi: 10.1111/ecin.12375.

[7] A. Stenberg, “What does Innovation mean-a term without a clear definition,” 2017.

[8] B. Wilkerson and L. K. L. Trellevik, “Sustainability-oriented innovation: Improving problem definition through combined design thinking and systems mapping approaches,” *Thinking Skills and Creativity*, vol. 42, Dec. 2021, doi: 10.1016/j.tsc.2021.100932.

[9] C. Hoolohan and A. L. Browne, “Design thinking for practice-based intervention: Co-producing the change points toolkit to unlock (un)sustainable practices,” *Design Studies*, vol. 67, pp. 102–132, Mar. 2020, doi: 10.1016/j.destud.2019.12.002.

[10] F. Pacheco, C. Rangel, J. Aguilar, M. Cerrada, and J. Altamiranda, “Methodological framework for data processing based on the Data Science paradigm,” Nov. 2014. doi: 10.1109/CLEI.2014.6965184.

[11] J. Aguilar, O. Buendia, A. Pinto, and J. Gutiérrez, “Social learning analytics for determining learning styles in a smart classroom,” *Interactive Learning Environments*, vol. 30, no. 2, pp. 245–261, 2022, doi: 10.1080/10494820.2019.1651745.

[12] M. Sánchez, J. Aguilar, J. Cordero, P. Valdiviezo-Díaz, L. Barba-Guamán, L. Chamba-Eras. Cloud “Computing in Smart Educational Environments: Application in Learning Analytics as Service”. In: Rocha, Á., Correia, A., Adeli, H., Reis, L., Mendonça Teixeira, M. (eds) *New Advances in Information Systems and Technologies*. Advances in Intelligent Systems and Computing, vol 444, pp 993–1002, 2016, https://doi.org/10.1007/978-3-319-31232-3_94.

[13] A. Gutiérrez, J. Aguilar, A. Ortega, and E. Montoya, “Autonomous cycles of data analysis tasks for innovation processes in MSMEs,” *Applied Computing and Informatics*, 2022.

[14] C. J. Hutto and E. Gilbert, “VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text,” Eighth International Conference on Weblogs and Social Media (ICWSM-14), 2014.

[15] A. Kumar and R. Jain, “Sentiment analysis and Feedback Evaluation,” in *Proceedings of the 2015 IEEE 3rd International Conference on MOOCs, Innovation and Technology in Education, MITE 2015*, Jan. 2016, pp. 433–436. doi: 10.1109/MITE.2015.7375359.

[16] O. Terán and J. Aguilar "Modelo del Proceso de Influencia de los Medios de Comunicación Social en la Opinión Pública", *Educere*, vol. 22, no. 71, 2018, pp.179-191. https://www.redalyc.org/articulo.oa?id=35656002014

[17] H. Sánchez, J. Aguilar, O. Terán, J. Gutiérrez de Mesa, Modeling the process of shaping the public opinion through Multilevel Fuzzy Cognitive Maps, Applied Soft Computing, vol. 85, 2019, https://doi.org/10.1016/j.asoc.2019.105756.

[18] J. Aguilar, “Different dynamic causal relationship approaches for cognitive maps”, Applied Soft Computing, vol. 13, pp 271-282, 2013, https://doi.org/10.1016/j.asoc.2012.08.037.

[19] J. Aguilar, “A Fuzzy Cognitive Map Based on the Random Neural Model”. Lecture Notes in Computer Science, vol 2070, pp. 333–3382001. https://doi.org/10.1007/3-540-45517-5_37

[20] Khramoin I.V, Kamenskaya M.A, and Tikhomirov I.A, *Sentiment Analysis of Innovations in Russian Media*. Proceedings of the 2nd Russian-Pacific Conference on Computer Technology and Applications, 2017.

[21] N. Haefner, J. Wincent, V. Parida, and O. Gassmann, “Artificial intelligence and innovation management: A review, framework, and research agenda☆,” *Technological Forecasting and Social Change*, vol. 162, Jan. 2021, doi: 10.1016/j.techfore.2020.120392.

[22] F. Zhou, R. J. Jiao, and J. S. Linsey, “Latent customer needs elicitation by use case analogical reasoning from sentiment analysis of online product reviews,” *Journal of Mechanical Design, Transactions of the ASME*, vol. 137, no. 7, Jul. 2015, doi: 10.1115/1.4030159.

[23] J. Füller, K. Hutter, J. Wahl, V. Bilgram, and Z. Tekic, “How AI revolutionizes innovation management – Perceptions and implementation preferences of AI-based innovators,” *Technological Forecasting and Social Change*, vol. 178, May 2022, doi: 10.1016/j.techfore.2022.121598.

[24] C. Kakatkar, V. Bilgram, and J. Füller, “Innovation analytics: Leveraging artificial intelligence in the innovation process,” *Business Horizons*, vol. 63, no. 2, pp. 171–181, Mar. 2020, doi: 10.1016/j.bushor.2019.10.006.

[25] T.-T. Kuo and N. Peng, “Exploiting Latent Information to Predict Diffusions of Novel Topics on Social Networks,” 2012.

[26] A. Gutiérrez, J. Aguilar, A. Ortega, and E. Montoya, “Intelligent systems in the innovation processes of textile MSMES-a Systematic Literature Review,” *International Journal of Entrepreneurship*, vol. 25, no. 5, 2021.

[27] J. Aguilar, A. Garcès-Jimènez, N. Gallego-Salvador, J. De Mesa, J. Gomez-Pulido, À. Garcìa-Tejedor, "Autonomic Management Architecture for Multi-HVAC Systems in Smart Buildings," *IEEE Access*, vol. 7, pp. 123402-123415, 2019, doi: 10.1109/ACCESS.2019.2937639.

[28] J. Vizcarrondo, J. Aguilar, E. Exposito and A. Subias, "MAPE-K as a service-oriented architecture," *IEEE Latin America Transactions*, vol. 15, no. 6, pp. 1163-1175, 2017, doi: 10.1109/TLA.2017.7932705.

[29] R. Fornieles Sánchez, “De Lasswell a Gorgias: los orígenes de un paradigma,” *Estudios sobre el Mensaje Periodístico*, vol. 18, no. 2, Dec. 2012, doi: 10.5209/rev_esmp.2012.v18.n2.41043.

[30] G. v. Oliveira, F. P. Coutinho, R. J. G. B. Campello, and M. C. Naldi, “Improving k-means through distributed scalable metaheuristics,” *Neurocomputing*, vol. 246, pp. 45–57, Jul. 2017, doi: 10.1016/j.neucom.2016.07.074.

[31] J. Bughin *et al.*, “Artificial Intelligence the Next Digital Frontier?” [Online]. Available: www.mckinsey.com/mgi.

Appendix D

Definition of innovation problems in organizations using Data Analysis from Hybrid Sources: Social Networks and Organizational Databases

Definition of innovation problems in organizations using Data Analysis Tasks from Hybrid Sources: Social Networks and Organizational Databases

Ana Gutiérrez[1,2,5], Jose Aguilar[2,3,4,5], Ana Ortega[2], Edwin Montoya[2,5]

[1] GIA, Universidad Francisco de Paula Santander, Av. Gran Colombia, Cúcuta, Colombia

[2] GIDITICS, Universidad EAFIT, Medellín, Colombia

[3] CEMISID, Universidad de Los Andes, Núcleo La Hechicera, Mérida, Venezuela

[4] IMDEA Networks Institute, Leganés, Madrid, Spain

[5] Tepuy R+D Group. Artificial Intelligence Software Development. Mérida, Venezuela

Abstract

Nowadays, we can express the experience we have lived with the products we use. Most of the time, we interact with brands and let our likes and dislikes be seen on digital platforms, either by interacting with social networks, filling out satisfaction surveys, or registering requests, complaints, and claims. To get the value of all available customer data of the companies, in this article, we propose to study the data from different sources using an autonomic cycle of data analysis tasks to define innovation problems in an organization. The tasks of the autonomic cycle filter the customer comments from different sources (e.g., from social networks, PCCS (petitions, complaints, claims, suggestions) systems of organizations, etc.), obtain their keywords, and analyze the patterns of the users to answer the questions of the 5W methodology (what, who, where, when and why?), in order to define innovation problems. Finally, this article analyzes a case study of a fashion company using its PCCS system and comments of Twitter, to identify useful information. Part of the information discovered was the reasons for customer returns, merchandise delivery problems, shipment failures, failure to respond timely to customers, among other things. With this information, the autonomic cycle is able to define customer and organization-oriented innovation problems, in order to respond to these identified problems.

Keywords: Innovation Problems, Autonomic Computing, Data Analysis, Machine Learning,

1. Introduction

Data analysis from different sources can be useful to define innovation problems by providing detailed information about customer needs and wants, market behavior patterns, and industry trends [1][2]. Analysis of data collected from these sources can help identify common problems and improvement opportunities for innovation. In the business context, it has been used to investigate the factors that influence the innovative behavior of employees [3], to identify innovative strategies that retail companies can use to stay competitive [4][5], to improve processes and services as a factor of competitiveness [6], to study customer experience [7][8][9][10], all previous cases of paramount importance for managers to identifying opportunities for innovation and constant improve.

Managing innovation within organizations implies designing strategies and policies to develop new products, processes, and services [10], improving their ability to compete in a constantly evolving market [11], and generating strategies for the improvement of production [12]. Using big data analysis will allow companies to analyze large amounts of information to identify patterns and trends that can lead to the identification of new business opportunities and improve decision making [13].

On the other hand, understanding social media data is important for defining innovation problems, as it provides valuable information about customer interests, behaviors and needs, the problems they face, opportunities for product and service improvements, among other things. This data can be used to develop innovative solutions that better meet customer needs [2][4][15]. They can also be used to identify strategies to improve business commercialization and marketing [16]. Another source of data that can be used is the PCCS system since they provide valuable information about problems and opportunities for improvement in the products, services and processes of a company. Therefore, these different data sources can be used to improve customer service, customer satisfaction, and brand reputation, to identify market trends to help develop innovative solutions [17], and to analyze customer complaints to improve services [18], improve the quality of services and identify factors that lead to low levels of customer satisfaction [19]. Thus, these different data sources can help companies to have additional sources of information to identify and define innovation issues.

This paper presents an autonomous cycle of data analysis tasks (ACODAT) for defining innovation problems based on sarcasm analysis in Satisfaction Rating surveys, and customer comments in social networks and PCCS systems. In preliminary works, we have defined the concept of ACODAT [20] [26] [27] [28], designed generic ACODATs for the management of innovation processes in organizations [1], and defined sentiment analysis approaches for the definition of innovation problems [2]. In this article, with the aim of improving the definition of innovation problems, an ACODAT is designed to study sarcasm in customer comments in PCCS systems, surveys and social networks. The main contributions proposed in this paper are:

1. Design an ACODAT to define innovation problems based on cynical or sarcasm analysis.

2. Define an approach to use different data sources with customer comments, in particular, in PCCS systems, surveys and social networks, for the definition of innovation problems.

3. Describe a case study of this ACODAT in the fashion textile industry.

The organization of this work is as follows. Section II describes related works. Section III describes the theoretical framework of this study, in particular, the ACODAT for defining innovation problems. Section IV describes the instantiation of this ACODAT to analyze customer comments from PCCS systems, surveys and social networks. Section V presents a textile case study, including experiments and analysis of results.

Section VI provides a comparison of our proposal with other studies, and finally, the conclusions and future works are presented in the last section.

2. Related works

In this section, we describe the main latest papers associated with our approach. First, we present the papers about the use of data analysis to determine innovation problems [1][2] [10][13][14][21] [22]. Second, we present papers about data analysis from hybrid sources [12][6] [23].

2.1 Data analysis to determine innovation problems.

The article [1] of Gutiérrez et al. defines several ACODATs to manage innovation processes in a Micro, small and medium-sized enterprise (MSME), with their multidimensional data models, and the characterization of the AI techniques required in the case of the textile industry. The article [2] evaluates opinions from social networks, which are currently used to establish relationships with customers, to define innovation problems taking into account the feelings expressed in social networks. Jahan, et al. [13] study the impact that technology has had on business innovation. The authors analyze the role of Big Data attributes used in other studies for analyzing business efficiency from operational excellence, financial returns, and customer perspectives.

The paper [14] of Lendel et al. explores how data analysis techniques can identify and solve problems in innovation management. It presents a framework that combines quantitative and qualitative data analysis methods to analyze innovation problems and suggests strategies for addressing these problems. The research includes gathering information through analysis, questionnaires, and interviews, provides insights into the challenges faced in managing innovation processes, and offers recommendations for improving innovation management in companies. The work [21] uses analysis of data to determine problems in companies, in this case, linked to environmental impact. They develop strategies to reduce, reuse and recycle materials to cope with cultural, commercial, and technological changes. In this sense, they present a framework for the definition of innovative processes for material management, considering factors like data analysis, among others.

Lorduy et al. [10] emphasize the impact of innovation on national and regional administration. They analyze the national policies related to science, technology, and innovation from different sources (e.g., national organizations such as the observatories of science and technology, the national administrative departments of statistics, among others). Also, the researchers apply surveys to understand local business idiosyncrasies actors of the financial, public and university sectors, among others. In the article [22], the objective was to review the existing literature on data-driven innovation (DDI), to understand how companies can use data and analytical tools to facilitate innovation processes. The results showed various aspects, such as innovation models in the market, competitive advantages of the innovation, and the necessity of the democratization of data. However, they highlight that little has been studied on how data helps the innovation process, which would be valuable for companies to learn from their own failures and improve their innovation processes in the future.

2.2 data analysis from hybrid source

Zaman et al. [12] explored various algorithms for prediction with a combination of financial and social media data to predict stock market behavior. Financial data includes historical stock price information and economic data, while social media data captures users' opinions and sentiments about stocks and companies. This Integrated data analysis enabled the identification of relevant patterns about market trends, and most of the classifiers demonstrated an improved accuracy. Babu et al.[23] examined the impact of big data analytics on manufacturing and service organizations, as well as the ability to drive results and innovation in these companies. They applied data analytics to enhance innovation, explore new products, as well as address intellectual, legal and privacy issues, among other things, which can be shared across the industry to facilitate the innovation processes.

Wang, et al. [24] explored the use of multi-source data for the management of public health emergencies in China. They used the response scenario method, including risk analysis and operational plans at the micro level, as well as strategic decision making at the macro level. They applied the study for monitoring and response to the Ebola virus in China. It was evidenced that it is possible to detect and respond to public health emergencies at very early stages by analyzing data from different sources. In the article [25], Koukaras et al. developed a model to predict stock movement using Twitter and StockTwits data. They collected tweets from these platforms and financial data from Finance Yahoo, and implemented several classification models. The main innovation of this work is the integration of multiple sources to enhance stock prediction accuracy.

Finally, the article [30] focuses on evaluating sustainable development by analyzing normalized economic losses and changes in social structure. They utilize different sources, such as earth observation data, along with data from the national oceanic and atmospheric administrations, to examine the effects of floods, landscape changes, and climate change patterns. By integrating satellite-based with social media data from Twitter, they assessed the sentiment of affected individuals and measure the mental impacts of Hurricane Ida. The study demonstrates that combining of sources provides a more comprehensive understanding of the aftermath of the hurricane (they help understand the mental state of the community, identify affected populations and areas, and model the prevalence of epidemics).

According to the previous works in this section, there are articles related to the use of data analysis to determine innovation problems [1][2] and the use of Big Data to study business innovation and innovation capabilities [13][14]. On the other hand, there are data analysis works from hybrid sources, for example, using organizational data [23], social networks [24][12][25][30], images [30] or financial data [12][25]. They have been applied in different sectors such as manufacturing, health, and finance. This work includes again the concept ACODAT [1][20] to define innovation problems but with the difference that includes different data sources to improve the definition of the innovation problems. Thus, it discovers valuable information about the company's products, services and processes using satisfaction surveys and customers' opinions in social networks and the PCCS systems of the organization, to identify innovation problems.

3. Autonomic Cycle to analyze innovation problems

In this work, we use the ACODAT described in [1]. In that work, the authors defined an ACODAT for the definition of innovation problems, and in [2], the authors instanced this ACODAT using social networks. They apply sentiment analysis on negative Twitter to build innovation problem statements. For the definition of the ACODAT is used the 5W approach (see Figure 1) [30].

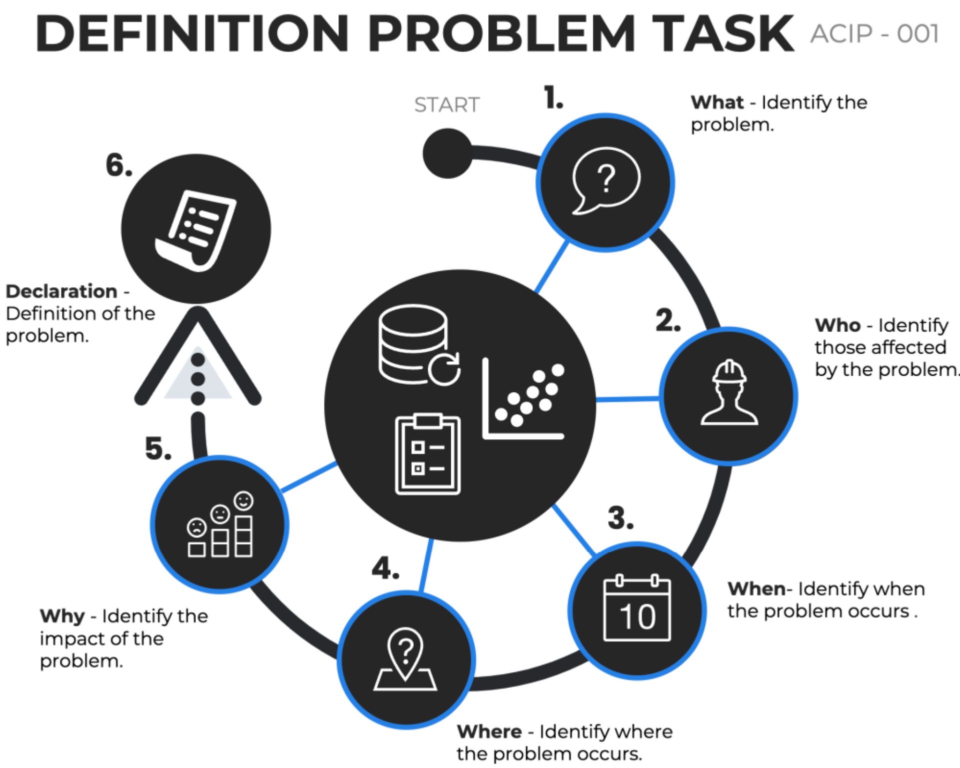

Figure 1. Structure of the ACODAT for the definition of innovation problems [1].

Each of the tasks of this ACODAT is described below:

1. What: In this task, innovation problems are detected. The objective of the task is to determine the innovation problems. For this case, descriptive and detection models can be used.

2. Who: this task identifies those individuals who are impacted by the innovation problems (e.g. customers, organizations, or specific groups). It makes use of descriptive models.

3. When: this task determines the time when the issue will arise, and it uses detection or prediction models.

4. Where: in this task, the location of the issue is determined using models for diagnosis.

5. Why: in this task, the problem's reasons are determined.

6. Declaration: The innovation problems are defined by this task. This task describes the statement of the innovation problems using Natural language processing (NLP) techniques.

4. Instantiation of the ACODAT for the definition of innovation problems using hybrid data sources

In this section is defined the instantiation of the ACODAT utilizing different data sources. In this case, we use PCCS systems, social networks, and some market studies based on customer satisfaction interviews. The macro-algorithm in Figure 2 defines the instantiation of the ACODAT for the "Definition of the Innovation Problem" based on [1][2], which is the basis of our SAHDS system (Sarcastic Analysis from Hybrid Data Sources) system.

Input: Dataset

Output: Definition of Problem

4. Analysis of the data sources to Identify sarcasm
    1. Filter the data sources
    2. Preprocess information
    3. Standardize information
    4. Apply Sentiment Analysis:

        Social Network -> Negative Comments

        Surveys -> Text with Positive Sarcasm

Figure 2. SAHDS-IPO macro-algorithm

The steps of this macro-algorithm are detailed next:

1. For the first task/step, several sub-steps are required:

   a. Select the sources of the information to be used (surveys, social networks, etc.)

   b. Prepare the data for the analysis. Specifically, NLP operations are performed, including tokenization, clean character and stopword removal, and stemming/lemmatization.

   c. Normalize the data.

   d. Analyze the sentiment of the text. Then, we check if the polarity is negative and the overall sentiment is negative, suggesting that the text might be sarcastic [31][32][33].

   e. Sort the data to utilize the most pertinent information for our scenario (in our case, only sarcastic positive cases).

2. Perform a cluster analysis based on the 5W model responses.

3. Create an innovation problem statement using the results of previous tasks.

Table 2 shows the modification of the ACODAT for the definition of innovation problems based on the SAHDS-IPO macro algorithm in Fig. 2.

Autonomous Cycle Task	**Macro-algorithm**		**Task improvement**	**Data Source**	
	Social Network	**Database**		**Social Network**	**Database**
Task 1. Identify sarcasm	1.a. Preprocess informa 1.b. Standardize informa 1.c. Identify Keywords 1.d. Determine sarcasm	1.a. Preprocess informat 1.b. Standardize informa 1.c Identify Keywords 1.d Determine sarcasm		Twitter	Satisfaction Rating surveys Customer Comments on PCCS systems
Task 2. What	2. Clustering based on Keywords	2.a Clustering analysis using the customer comments	The comment of the customer is used to identify what is the problem	Twitter	Satisfaction Rating surveys
		2.b Clustering analysis ba the typification of the pr in the PCCS systems	The typification (complaint/claim) will allow identifying the reason for the problem		Analysis of PCCS systems
		2.c Clustering analysis ba answer-request on market study	Identify if the answer gives a solution to the request		Market Study: Answer_according request Y-N Again_talk_Virtual-Asistent
		2.d Clustering analysis ba deficiencies	Identify deficiencies in the product as variety, quality, difficulty finding the product		Product : Quality Variety KidnessAtention EasyFindProduct OffeerCompletLook

Task 3. Who	3. Clustering based on id_users	3.a Clustering analysis based on ids in the PCCS systems 3.b Clustering analysis ba assigned agent	Identify the external customer affected by the problem and the internal customer who will solve the problem.	Id users	PCCS system: Id
Task 4. When	4.Clustering analysis bas on the date	4.a Clustering analysis ba on the date	Identify the date: PCCS system or survey	Date, month, working day of the Twitter	PCCS system: date Satisfaction Survey: date
Task 5. Where	5.Clustering analysis bas on localization	5.a Clustering analysis ba unity 5.b Clustering analysis ba origin	Identify the unit responsible for solving the PCCS request or the place where answering the survey	Localization	PCCS system: Responsible_unity
Task 6. Why	6.Clustering analysis bas on negative sentiment	6.a Clustering analysis based on response times	Identify the waiting time for a response to a PCCS	Compound	PCCS: response times (1-40 days)
		6.b Clustering analysis ba satisfaction rating	Identify the level bad of satisfaction		Satisfaction Rating: The level of satisfaction
		6.c Clustering analysis ba recommend brand	Identify how highly would you recommend this brand to others.		Market Study: Recommend brand (1-10)
		6.d Clustering analysis b difficulty	Identify difficulty in handling the request		Market Study: Manage request
		6.e Clustering analysis b customer experience	Identify customer experience (good, bad)		Market Study: customer experience
		6.f Identify the toxicity le sarcasm of the comments	Identify communication level (easy or difficult)		Market Study: Easy_communication

		6.g Clustering analysis ba sarcasm identified using comments	Identify the sarcasm in the comments		Satisfaction Rating: Customer Comments
Task 7. Declaration	7. Definition of the Problem	7a. Definition of the Problem	Problem Declaration based on the previous answers	NLP	NLP

Table 2. ACODAT adapted to the SAHDS-IPO macro-algorithm

Each task of the adapted ACODAT is described below.

Task 1. Identify sarcasm. This task is defined by different processes to analyze the data sources:

Information preprocessing: In this step, the comments from the organizational databases or social networks are transformed. Specifically, the text is tokenized, stop words are removed, special characters are sanitized, and the text is stemmed/lemmatized. Figure 3 shows a sub-algorithm (a typical NLP technique) for preprocessing comments. With the PCCS data is not necessary to do this step, because the typification is normalized when it is predefined during the design of the PCCS system. The typification predefines the types of registers (request, complaint, claim, suggestion) of a client in the PCCS system.

Input: Text Comment

Output: Clean Text Comment

1. Input Text Comment
2. Remove Stop Words
3. Remove URLs
4. Remove special characters

Figure 3. Information preprocessing sub-algorithm

Standardize the information: this step normalizes the information to be used. Figure 4 shows the sub-algorithm that defines the steps to extract the information from the survey dataset. For example, the year is extracted from the date value in the survey and the PCCS systems. Similarly, it is done with the rest of the information of interest contained in the comments or PCCS systems to answer the questions of the 5W model (e.g., responsible to solve the problem, localization of the origin of the petition, etc.). In the case of the day, if the hour is less than noon (12) then the day column has a value of 1 (it means morning), else it will be 2 (it means afternoon).

Input: Dataset

Output: Standardize Dataset

1. Standardize Information

   a. Extract from the date field the value of the year

   b. Extract from the date field the value of the month

   c. In the survey, If rating <= 3:

      a. Include = 1

   else

      b. Include = 0

   d. In the social network

      If hour < 12:00pm : a. Working day = 1

      else b. Working day= 2

   e. Extract from the data the text

Figure 4. Standardize information sub-algorithm

Keyword acquisition: In this step, the comments are preprocessed to obtain the keywords. In the social network are used the user's comments, in the satisfaction surveys, the answers of the users, and in the case of the PCCS system, the classification established in the company for the comments. With this data is determined the keywords using the TF-IDF metric [36].

Apply Sentiment Analysis: According to the keywords previously identified by text, then we use vaderSentiment library [29] in order to identify the sarcasm. It's important to mention that Sarcasm is a form of figurative language, where the speaker expresses his/her thoughts in a sarcastic way [34]. This library defines word polarities and classifies them as positive, negative, neutral, or compound (hybrid sentiment). Additionally, this library defines a composite score value for each text. It represents the average polarity of all words in the text. Figure 5 shows the macro algorithm used to perform sarcasm analysis on the dataset. If the composite score is less than 0.1, then the sarcasm is positive, otherwise, the sarcasm is negative. The next step is to filter the comments. In the social network, the best keywords are the negative words, while in the surveys the keywords are the sarcasm words. The PCCS uses the typification. This new subset of sarcastic comments serves as a source of information for continuing the next task. By analyzing and categorizing the keywords, organizations can identify trends, problem areas and opportunities for improvement. This allows them to take concrete steps to address recurring problems, implement solutions and strengthen their relationship with customers.

```
Input: Initial Dataset

Output: Sarcasm Text

1. Apply sentiment analysis

If Compound_Score < -0,1

        a. Sarcasm=Positive

        else

        b. Sarcasm=Negative

If Negtive_Score < 0
```

Figure 5. Sarcasm Analysis sub-algorithm

Task 2. What: Identify the problem: This task requires the definition of groups that characterize the innovation problems. Figure 6 shows a sub-algorithm that defines the clustering problem based on the keywords obtained previously. For this task, the silhouette metric with the K-means clustering algorithm was used to determine the numerical value of k [29, 35]. Silhouette values typically range from -1 to 1, with high values indicating that the data fit its own cluster well. In general, silhouette values greater than 0.5 imply good clustering results [29, 35].

```
Input: Keywords of the previous phase

Output: Clusters

1. Call the Clustering technique
```

Figure 6. Keyword-driven Clustering sub-algorithm

Finally, this task analyzes the centroid of each cluster to determine the answer to the "what" of the 5W model using the keywords [30]. In summary, Figure 7 shows the main steps to determine the 'what'.

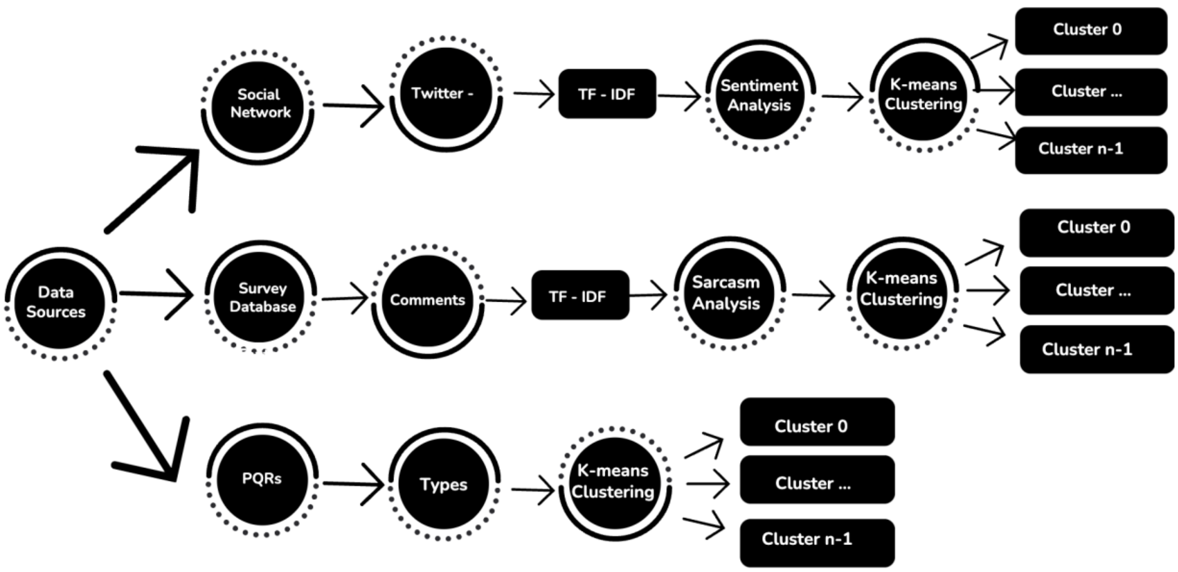

Figure 7. Determination of the "What" of the 5W model

Task 3. Who: Identify who is affected by the problem: this task uses the information from the previous task to identify who is affected by the problem. For this purpose, for each cluster obtained in Task 1, a new clustering is performed (See Figure 8) based on comments or aspects to be improved that the clients made in the surveys. In the PCCS system, it is also possible to identify who is affected by the problem. In this case, the entity or dependency that will resolve the petition, complaint, claim or request. For social networks, an external client is identified who writes the text or comment, or an internal client who responds to the comment. In this case, an individual may be an employee of the business, a physical or online consumer, or an entity of the business that participates in multiple requests by responding to different requests.

Input: Comments Clusters and user-Ids or unity responsible clusters

Output: Who

1. Call the Clustering technique based on

    a. User information from social network comments

    b. User or the responsible unity in the PCCS system

Figure 8. User-driven Clustering sub-algorithm

Task 4. When: Determine when the problem occurs: This task uses the information collected in task 1 to identify when the problem occurs. In this case, a new group is created for each cluster of the first task using the date as the similarity criterion (see Figure 9).

Input: Comments Clusters and date and PCCS Clusters date

Output: When

1. Call the Clustering technique based on date information

    a. Publication in the social network
    b. Completion of the survey

Figure 9. Date-driven Clustering sub-algorithm

Task 5. Where: Identify where the problem occurs: This task uses the information collected in task 1 to identify where the problem occurs. Again, in the social network, it is the place where the comment was published (geolocation). In other cases is used the customer's stores where they answer the survey. This information can be determined when they visit a physical store in the malls, or outlet stores, or maybe they call the phone line. In the PCCS system is the origin from where is registered the petition (see Figure 10).

Input: Stores and origin petition Clusters

Output: Where

1. Call the Clustering technique based on origin information

    a. Geolocation from the social network comment
    b. Place where is answered the survey

Figure 10. Site-based clustering sub-algorithm

Task 6. Why: Identify the impact of the problem: This task uses the information collected in task 1 to identify the impact of the problem. In the social network is used the compound score value of the task of sentiment analysis to determine the impact of the problem. According to this value, the impact can be low, medium or high. In the survey is based on questions about satisfaction on the products, experience, quality, kindness in the attention, etc. (the scores in these aspects). In the other case, we use the answer time in the PCCS system, which indicates whether the stores are effective in meeting the customer's request (see Figure 11).

Input: Questions Clusters and answer time PCCs

Output: Why

1. 1. Get compound_score value from social network
   if compound_score is between [-0,01 , -0,33) then

   a. Why = High

   else if compound_score is between [-0,33 , -0,66):

   b. Why = Medium

   else c. Why = Low

2. Get answer value from the survey

   if the score is between [1 , 3):

   a. Why = High

   else if the score is between [4, 6):

   b. Why = Medium

   else c. Why = Low

2. Get time value PCCs

Figure 11. Impact Identification sub-algorithm

Task 7. Declaration: Definition problem: In this task, we use all the results of the previous tasks to define the innovation problem. Specifically, in this work, several sentences are created, one oriented to the Petitions, Complaints and Claims of the clients, and others taking into account the comments where the clients expressed their thoughts in the surveys or social networks. Below are defined these sentences of the innovation problems:

Sentence -> using Social Networks:

<Who(user)> + "published" + <When(date)> + "from" + <Where(localization)> + "about" + <What(keywords)> + "such that the impact is" + <why><why(compoundSa)>

Sentence -> using the PCCS system of the organization:

<Who(assigned agent_pcc)> + "received" + <When(date_pcc)> + "request from" + <Where(Ticket)> + "about" + <What(typifaction_pcc)> + "such that the impact is" + <why(time_pcc)>

Sentence -> using Customer opinions in the surveys:

<Who(id_cliente> + "published" + <When(date survey)> + "from" + <Where(SurveyName)> + "about" + <What(list of aspects)> or <What(answer-request)> or <What(deficiencies)> + "such that the impact is" <Why(The level of satisfactionSurvey)> or <why(RecommendBrandSurvey)> or <Why(difficulty)> or <Why(customer experience_survey)> or <Why(communication)>

These statements can end up in a single statement that adds the results of all.

5. Experiment and Analysis of Results

This section presents the experimental context for the instantiation of the ACODAT for the Definition of the Innovation Problem using different data sources (social networks, surveys, and a PCCS system).

5.1 Case Study

In this study, we used datasets from the OFFCORSS company, which provided datasets of surveys applied to customers who responded in physical stores or outlets or made queries or requests through the virtual VoiceBot. We were also provided with a database containing the records of the PCCS registered by customers. Finally, the information from the company's Twitter account was used to determine the customers who followed it and gave their opinions about it. Table 3 summarizes the information used of each source.

Data Source	Data
Social Network	Twitter Comments, id_user, localization, date
Database Surveys	Satisfaction Rating: Customer Comments, Answer_according request Y-N, Again_talk_Virtual-Asistent, Quality, Variety, KidnessAtention, id, date, EasyFindProduct, OffeerCompletLook, The level of satisfaction
PCCS system	Typification, id, date, Responsible_unity, response time, TicketOrigin

Table 3. Data sources and data

a. ACODAT in the case study

This section describes each task of the ACODAT for defining innovation problems.

Task 1. Sarcasm analysis. For the task, the first step is the PLN techniques to determine the information to process (in our case, the keywords). An example of keywords obtained from the different sources are

Data Source	**Keywords**
Social Network	Assistance, robbery, vigilance
Surveys	Suggestions, ¨Need help¨
PCCS system	TV-devolution, TV-Return Request, VD-Data update, VD-Payment Verification,

Table 4. Example of keywords by source

For the surveys, we take the keywords of the comments and apply sentimental analysis using vaderSentiment Library. Then, we get the value for each text according to the compound value, and particularly, the sarcasm is true when the value is lower than -0.1. Table 5 presents a partial view of the results of the sarcasm analysis for this dataset.

ITEMID	ANSWERVALUE	ORIGEN	compound_Score	Negative_Score	Positive_Score	Neutro_Score	sarcasmo
200	1no resolvieron Nada. No es comprensible lo qu...	37	-0.2960	0.155	0.0	0.845	True
200	Hola	37	0.0000	0.000	0.0	1.000	False
200	Excelente!	37	0.0000	0.000	0.0	1.000	False
200	Hola	37	0.0000	0.000	0.0	1.000	False
200	Fosl	37	0.0000	0.000	0.0	1.000	False
...	...	...	...	...	...	...	...
616	Disponibilidad de zapatos de tallas 24	54	0.0000	0.000	0.0	1.000	False
616	El stock para todas las tallas, es decir si se...	54	-0.2960	0.073	0.0	0.927	True
200	Necesito que me ayuden con un paz y salvo y no...	37	-0.5267	0.196	0.0	0.804	True
200	buenas tardes,_x000D_\n_x000D_\naun no tenido ...	37	-0.5267	0.112	0.0	0.888	True
200	Necesito completar mi compra, pero aun no teng...	37	-0.2960	0.115	0.0	0.885	True

Table 5. Results of the Sarcasm Analysis step

In the PCCS system, the predefined typification is used. Table 6 presents some of the predefined examples of PCCS typification’s.

	ID	State	Affair	Responsible	Required	Month	Date	Assigned	Update	Typification
0	746861	Cerrado	TV-Eventos promocionales	Centro De Experiencia	2021-12-06	Diciembre	2021-12-07	Servicio al cliente TV	2022-01-08	TV-Eventos promomocionales
1	748684	Cerrado	TV-CancelaciÃ³n de Pedido	Centro De Experiencia	2021-12-11	Diciembre	2022-01-03	Servicio al cliente TV	2022-01-08	TV-Cancelación de pedido
2	749790	Cerrado	TV-CancelaciÃ³n de Pedido ADDI	Centro De Experiencia	2021-12-15	Diciembre	2021-12-21	Servicio al cliente TV	2022-01-08	TV-Cancelación de pedido
3	750162	Cerrado	TV-Solicitud DevoluciÃ³n AutomÃ¡tica	Centro De Experiencia	2021-12-16	Diciembre	2021-12-24	Ventas Tienda Virtual	2022-01-08	TV-Solicitud Devolución
4	751229	Cerrado	TV-Solicitud DevoluciÃ³n AutomÃ¡tica 2*1	Centro De Experiencia	2021-12-20	Diciembre	2021-12-29	Ventas Tienda Virtual	2022-01-08	TV-Solicitud Devolución

Table 6. Results of examples of PCCS typification's.

The social network uses negative Twitter comments. In this step, for sentiment analysis is used the vaderSentiment library in each keyword in the Twitter register (column of text) to determine each person's mood. This library calculates the Compound_score of each text. Table 7 provides a partial view of the sentiment analysis results for this dataset. Note that tweets are negative when Compound_score < 0.

TweetE	compound_Score	Negative_Score	Positive_Score	Neutro_Score
@DanielSamperO @fdbedout there is no gentleman...	-0.2960	0.115	0.000	0.885
RT @Dantestereo Putting the former prisoner in...	-0.9231	0.425	0.000	0.575
Putting the ex-prison in prison for bribing wi...	-0.8689	0.380	0.000	0.620
@OFFCORSS Perhaps if today in the world and in...	-0.9074	0.197	0.046	0.757

Table 7. Results of the Sentiment Analysis for Twitter

Task 1. What: Identify the problem: This task defines clusters that identify innovation issues related to comments in the surveys, social networks or PCCS systems. The keywords in the comments or typifications are used by the K-means clustering algorithm. In the case of surveys, five clusters were found according to the values of the silhouette metric (see Table 8).

N clusters	Average silhouette
0	0.70
1	0.72
2	0.75
3	0.80
4	0.87

Table 8. Clusters based on the keywords in the case of surveys.

The cluster centroids defined by this task are shown in Table 9. It shows the keywords for each cluster that should be used to define the innovation problem.

Cluster centroids				
Cluster 0:	**Cluster 1:**	**Cluster 2:**	**Cluster 3:**	**Cluster 4:**
super	order	does	dont	did
suggestions	received	solution	answer	solve
moment	address	request	improve	concern
brand	able	payment	sizes	answer
boxes	resolved	attention	think	thank
aisles	time	good	sent	help

Table 9. Keywords of the centroid of each cluster for the surveys

For social network has been used the same process explained in the article [2]. Thus, we use the keywords in negative tweets to apply the K-means clustering algorithm. The results are shown in Table 10.

Cluster centroids			
Cluster 0:	**Cluster 1:**	**Cluster 2:**	**Cluster 3:**
hotline	negative	worst	terrible
stolen	violated	answer	service
minutes	makes	happened	bought
waiting	responding	email	bad
follow	rating	assistance	refund
resources	platform	robbery	requested
gentleman	timely	say	claim
size	habeas	vigilance	WhatsApp
small	report	bonus	requesting
price	data	consumer	process

Table 10. Keywords of the centroid for some clusters for the social network

Finally, for the PCCS system is used its typification. Some examples of typification are TV-Return Request, VD-Data update, VD-Payment Verification, etc. The typification identifies the main words that describe the problem (see in Table 11 some clusters).

Cluster centroids		
Cluster 0	**Cluster 1**	**Cluster 2**
TV-Return Request	VD-Payment verification	VD-Incentive validation query
TV-Order status	VD-Modification of quota	VD-Affiliation status
VD-Data update	VD-Request for payment	VD-Right of request
TV-Order with new delivery	VD-Request for Order Cancellation	VD-Information on how to place an order via WEB

Table 11. Centroids for some clusters for the PCCS system

In the next task, a grouping process using different similarity criteria (user, date, location) is performed for each cluster obtained in this phase.

Task 2. Who: Identify the ones affected by the problem: In this task, we can observe who were affected by the problem. In the case of social networks, they can be external customers who post on the network, or employees who respond to customer comments. In the case of the surveys, they were always customers who visited the store or made a purchase in the physical or virtual store. In the case of the PCCS, it was possible to identify the area to which the request was assigned in order to provide a response and solution to the customer who submitted the request. In this case, the units responsible for the PCCS were standardized. The assigned units were as follows: 1 -> Experience Center, 2 -> Virtual Store Sales, 3 -> CTS Transportation, 4 -> Portfolio, 5 -> Returns Area, 6 -> Customer Service TV, 7 -> Customer Service, 8 -> Legal, 9 -> Treasury, 10 -> Affiliation Area, 11 -> Administrators, 12 -> Marketing and Sales (See Table 12).

Data Source	**Identify Who**	**Clusters**
Social Network	Clothing customers	External users or internal users
Surveys	Clothing customers	Physical store customers, Online store customers, or Outlets customers
PCCS system	Units responsible	Experience Center, Virtual Store Sales, CTS Transportation, Returns Area, Customer Service TV

Table 12. Identify Who

Task 3. When: Identify when the problem occurs. In this task, it became evident that the PCCS was filed in the months of November and December 2021, and the responses to the surveys answered by customers were carried out in 2021. In the case of social networks, the Twitters were in July and August 2021.

Task 4. Where: Identify where the problem occurs. In this task was used the place where they publish (see Table 13). In the case of social networks, geolocation was used in order to identify the city of the publication. In the case of surveys, directions of stores and the virtual Voicet IP were used to identify the place where customers were located. For the PCCS, the place was taken from the value indicated by the system where the request was received.

Data Source	**Identify Where**
Social Network	Geolocation
Surveys	Virtual Voicet IP, Store location
PCCS system	Website or Chat IP, or social network used to connect to the PCCS system

Table 13. Identify Where

Task 5. Why: Identify the effect of the problem. This task took into account the information on the time used to respond to the PCCS, representing this measure the impact that receiving a timely or delayed response has for the client according to the time spent by the unit provided. Some results for time are 0 days, or 7 seven days, with 0 days, implies that a timely response was given, and the impact is low. While the responses took between 0 and 7 days, which implies a medium impact, considering that the organization has up to 15 days to respond to the PCCSs received. For the answers analyzed by the surveys, the sensations received from 0 to 3 were taken into account for some of the following questions: about satisfaction in general, of the products, experience, variety, quality, kindness in the attention. Finally, in the case of social networks, a compound value from (-0.66 to -1] corresponds to a weak impact, from (-0.33 to -0.66] to a medium impact, and from (-0.01 to -0.33) to a strong impact.

Task 6. Declaration: Definition of the problem: This task uses all the results of the previous task to define innovation problems. According to the patterns defined in section IV, Table 14 shows the interpretation of the information for cluster 2 of the social network, Table 15 shows the Interpretation of the information for cluster 3 of the surveys, and Table 16 shows the interpretation of cluster 0 of PCCS.

b. Discussion of the results

This section presents the results for each data source (social network, PCCS system, and surveys).

1. Results for cluster 2 of social networks.

Task.		Results
8.	What	The worst answer, happened, robbery, say.
9.	Who	Clothing customers and one internal user
10.	When	In the month of July 2021 in the afternoon session.
11.	Where	The publications are from Tunja, Medellin, and Bogotá.
12.	Why	The impact is Medium

Table 14. Interpretation of the information for cluster 2 of the social network

Problem definition:

Using the pattern of the Sentence for social networks, the statement of the innovation problem is:

Clothing customers ***published*** *in July 2021 in the afternoon session* ***from*** Tunja, Medellin, and Bogotá and ***about*** *Worst answer, happened, robbery, say.* ***The impact is*** *Medium.*

In accordance with the previous clause, non-compliance in the delivery of goods, as well as purchases that are returned or do not come, may constitute a violation of consumer rights. This suggests that in order to enhance customer service in general, innovation in the logistics and transportation process must be considered. For example, intelligent fleets can be designed with the use of sensors and data analysis to optimize routes, delivery times and vehicle maintenance. It is also possible to explore the use of drones to deliver packages in areas of difficult access or high traffic congestion.

2. Results for cluster 3 of the Customer survey

Task.	Results
1. What	*don't, answer, improve, size, think, sent*
2. Who	Customers
3. When	Year 2021
4. Where	Outlets Stores
5. Why	Acceptable level with respect to the variety of products in the store, of the other variables there are no ratings less than or equal to three.

Table 15. Interpretation of the information for cluster 3 of the customer survey

Problem definition:

Using the pattern of the sentence for customer surveys, the statement of the innovation problem is:

Customers ***published*** *in the year 2021* ***from*** *Outlets Stores*, **about** *don't, answer, improve, size, think, sent* **such that the impact is** *acceptable.*

In accordance with the previous clause, customers generally complain about lack of responsiveness, shipping failures, and lack of guidance and advice on marketing products, among other things. For example, innovation in product offering and size availability must be defined, so garments could be designed to be adjustable and adaptable. Thus, garments could have features such as adjustable elastic waistbands, variable length straps, or detachable parts. Innovation in garment sizing has the potential to transform the shopping experience and improve customer satisfaction by providing garments that fit more precisely and comfortably to a wide range of bodies.

3. Results for cluster 0 of the PCCS system

Task.	Results
1. What	VD- *Data update*
2. Who	*Affiliations area*
3. When	*November and December 2021*

4. Where	Zendesk
5. Why	0 – 7 days

Table 16. Interpretation of the information for cluster 0 of the PCCS

Problem definition:

Using the pattern of the sentence for PCCS systems, the statement of the innovation problem is:

The affiliations area **received** in November and December 2021 **requests from** Zendesk **about** VD- Data update, **such that the impact** is medium.

According to the previous statement, customers in general are placing requests to change data, update personal data, quantity of items in the order, product sizes. Improving data updating can be approached from different perspectives, from process optimization to the adoption of advanced technologies. Innovation in this process seeks to achieve faster, more accurate and efficient updates that contribute to successful operations and decision making.

Definition of the innovation problem integrating the previous definitions

Proposing an innovative project to solve the problem of customer services and the lack of timely response, leads to the search for solutions to answer the following question: What to do?, First create mechanisms that lead to effective and efficient communication with the customer to build trust through digital platforms accessible, improving customer service in terms of online shopping, returns, complaints, claims and payment methods.

Also, our innovation proposal must focus on three key areas to enhance customer experience: efficient data updating, revolutionary garment sizing, and smart logistics management. By addressing these challenges, we aim to create a seamless, personalized, and reliable journey for customers, ultimately driving customer satisfaction, loyalty, and sustainable growth. Thus, our proposal must take advantage of the opportunity offered by the technologies by making changes in the marketing structure through an analysis of the external and internal environment of the company, reviewing and analyzing new strategies at technical, commercialization, and financial level, which must be frequently monitored and evaluated.

6. Comparison with previous works

In this section, we propose several criteria against which to compare our work with previous works. We define the following criteria:

Criterion 1: They use an ACODAT to define innovation problems.

Criterion 2: They use different data sources to define innovation problems.

Criterion 3: They manage innovation through Big Data in order to define innovation problems.

Table 17 shows the quality assessment of related works based on these criteria.

Work	Criterion 1	Criterion 2	Criterion 3
[1]	✓	✗	✗
[2]	✓	✗	✗
[6]	✗	✗	✗
[13]	✗	✗	✓
[22]	✗	✗	✓
This work	✓	✓	✓

Table 17. Comparison with preceding works.

For criterion 1, Gutiérrez et al. [1] define ACODAT for definite innovation problems and [2] carries the instantiation of this ACODAT. For criterion 2, Gutiérrez et al. [2] use social networks for the instantiation, but this work uses different data sources such as from social networks, surveys and the PCCS system, in order to define innovation problems. For Criterion 3, Sakila et. [13] and Samarasinghe et al. [22] manage innovation through Big Data to investigate the impact on innovation. Particularly, the article [2] discovers problems that motivate improving product and service innovation, while in this work, had the opportunity of discovering problems not only with products or services, but also oriented to improve processes, to think about innovations in marketing and organizational level.

As shown in Table 17, current papers do not satisfy all three criteria. On the contrary, this work full all the criteria because it uses ACODAT and different sources of data to improve the quality of the results obtained. With this work was possible to generate innovation proposals for different problems, such as such purchases that are returned, shipment failures, etc. The integration of the different data sources gives a holistic vision of the problems that allow improving both the issue identification and the innovation proposal. The combination of data sources allows for cross-validation of information. When insights are consistent across multiple sources, it increases the confidence in the accuracy of the findings. Also, can reveal hidden patterns that may not be apparent when examining data in isolation. Patterns that emerge from the synergy of multiple sources can offer deeper insights into customer behaviors and preferences.

Innovation is crucial for SME growth. Integrating insights from various platforms helps SMEs identify innovation opportunities that are rooted in real customer needs, reducing the risk of developing solutions that miss the mark. SMEs can quickly identify emerging issues and trends, enabling them to adapt their strategies and offerings promptly to meet changing customer demands.

7. Conclusions

The main contribution was to be able to use different data sources to improve the definition of the innovation problems in MSMEs. In this case, we have used sentiment analysis to determine if a text is sarcastic or not, and depending on the result, we worked with those texts that had a positive score for sarcasm. The PCCS data sources were also used, which proved to create and improve the problem sentences and their grammatical structure. As a result, It is improved the quality of the innovation proposal based on the problem that was defined. It confirms that different data sources can help companies to have an additional source of information that can identify and define innovation problems.

Thus, improvements in innovation processes are possible using data analysis tasks. The use of different data sources can discover valuable information about the company's products, services, and process pain points and areas for improvement. The combination of diverse data sources brings a synergistic effect, offering deeper, more accurate, and more actionable insights than analyzing each source in isolation. This approach leverages the strengths of each data source to provide a more holistic understanding of customer sentiments and innovation challenges. Combining real-time data from social networks and online platforms with internal data sources like PCCS systems or surveys enables the organization to gain insights and identify emerging issues in real-time. This proactive approach facilitates swift problem resolution.

The main limitation in the development of this work was that the databases used were at the local level, with the information provided by the studied organization, which meant using static data in a date range. Another limitation is the incomplete datasets, not all customers express their opinions online, which could lead to an incomplete representation of customer sentiments. This can skew the analysis and conclusions drawn. This work used data from different sources such as social networks, and PCCS systems. Thus, other limitations of relying on specific sources could include potential biases or incomplete representation of customer opinions if certain demographics or platforms are not well-covered. Also, data is constantly evolving. Thus, the clustering models need to adapt to new information, features, and trends to maintain their effectiveness over time.

One work in the future is the implementation of an online tool with a user interface that can be intuitive, easy to use and visually attractive to carry out the whole process of defining the problem with the use of these different tasks of data analysis. Another future work is to give continuity to the implementation of the stages of the innovation process, having as the source of data and inputs the results of this first sub-process of problem definition. The idea is to apply the ACODAT methodology to the other sub-processes in order to incorporate autonomous cycles of data analysis tasks. Finally, another future work is to implement the automatic formulation of the problem verifying if the formulation is appropriate, and correct, according to the context in which it is being used.

Reference

[1] Gutiérrez, A., Aguilar, J., Ortega, A. and Montoya, E. "Autonomous cycles of data analysis tasks for innovation processes in MSMEs", Applied Computing and Informatics, 2022, Vol. ahead-of-print No. ahead-of-print. https://doi.org/10.1108/ACI-02-2022-0048

[2] Gutiérrez, A., Aguilar, J., Ortega, A. and Montoya, E. “Sentiment analysis in the social networks for the definition of innovation problems in organizations.” Submitted to publication, 2022

[3] Fitriana, A. and Satrya, A. “Employee Service Innovative Behavior in the Public Sector: The Role of Proactive Personality and Work Engagement,” in Proceedings of the 19th International Symposium on Management (INSYMA 2022), 2023, pp. 423–432. https://doi.org/10.2991/978-94-6463-008-4_53.

[4] Bae, Y., Choi, J., Gantumur, M., Kim, N. Technology-Based Strategies for Online Secondhand Platforms Promoting Sustainable Retailing. Sustainability. 2022, Vol. 14. https://doi.org/10.3390/su14063259

[5] Luamba, D., Blye, D., Williams, D., and Chagadama, J. “Innovative Strategies for Small Retail Companies’ Sustainability,” International Journal of Business and Management Research, 2021, Vol. 9, no. 3, pp. 330–338. https://doi.org/10.37391/ijbmr.090311.

[6] Saldarriaga, M., Guzmán, M., and Concha, E. “Innovación Empresarial: Factor de competitividad y calidad de vida en Popayán, Colombia”. Revista Venezolana de Gerencia, 2019, Vol. 2, pp. 151-166. ISSN: 1315-9984.

[7] Hopkins, J. “An investigation into emerging industry 4.0 technologies as drivers of supply chain innovation in Australia,” Computers in Industry, 2021, Vol. 125, https://doi.org/10.1016/j.compind.2020.103323.

[8] Dekimpe, M., “Retailing and retailing research in the age of big data analytics,” International Journal of Research in Marketing, 2020, Vol. 37, no. 1, pp. 3–14, https://doi.org/10.1016/j.ijresmar.2019.09.001.

[9] Trivedi, J. “Examining the customer experience of using banking chatbots and its impact on brand love: The moderating role of perceived risk”. Journal of internet Commerce, 2019, Vol. 18, no 1, p. 91-111, https://doi.org/10.1080/15332861.2019.1567188

[10] Lorduy, O., and Rangel, J. “Análisis de indicadores de innovación para una muestra de empresas manufactureras de Montería (Colombia),” Semestre Económico, 2019, Vol. 22, no. 52, pp. 49–73, https://doi.org/10.22395/seec.v22n52a3.

[11] Shayah, H., and S. Zehou, “Organizational culture and innovation: A literature review”, in Proceedings of the 2019 3rd International Conference on Education, Culture and Social Development (ICECSD 2019), 2019, https://doi.org/10.2991/icecsd-19.2019.58.

[12] Zaman, N., Ghazanfar, M., Anwar, M., Lee, S., Qazi, N., Karimi, A., et al. “Stock market prediction based on machine learning and social sentiment analysis”. TechRxiv. Preprint. https://doi.org/10.36227/techrxiv.22315069.v1

[13] Sakila, J., and Mesbaul, S., “Innovation Management: Is Big Data Necessarily Better Data?”, Management of Sustainable Development, 2022, Vol. 14, no. 2, pp. 27–33. https://doi.org/10.54989/msd-2022-0013

[14] Lendel, V., Hittmar, S., and Siantova, E., “Identification of the Main Problems in the Management of Innovation Processes and the Draft of Appropriate Recommendations”, Lecture Notes in Business Information Processing, 2015, 10th International Conference, Maribor, Slovenia, https://doi.org/10.1007/978-3-319-21009-4_17

[15] Zhou, F., Jiao, R., and Linsey, J., “Latent customer needs elicitation by use case analogical reasoning from sentiment analysis of online product reviews”, Journal of Mechanical Design, Transactions of the ASME, 2015, Vol. 137, no. 7, https://doi.org/10.1115/1.4030159

[16] Gutiérrez, A., Aguilar, J., Ortega, A. and Montoya, E, “Intelligent Systems in the Innovation Processes Of Textile Msmes-A Systematic Literature Review,” International Journal of Entrepreneurship, 2021, vol. 25, no. 5.

[17] Joung, J., Jung, K., Ko, S., and Kim, K., “Customer complaints analysis using text mining and outcome-driven innovation method for market-oriented product development,” Sustainability (Switzerland), 2019, Vol. 11, no. 1., https://doi.org/10.3390/su11010040

[18] Roper, S., and Bourke, J., “Innovating into trouble: When innovation leads to customer complaints,” Res Policy, 2022, Vol. 51, no. 10, https://doi.org/10.1016/j.respol.2022.104593

[19] Ghazzawi, A., and Alharbi, B., “Analysis of Customer Complaints Data using Data Mining Techniques”, Procedia Computer Science, 2019, Vol. 163, pp. 62-69, https://doi.org/10.1016/j.procs.2019.12.087

[20] Pacheco, F., Rangel, C., Aguilar, J., Cerrada, M., and Altamiranda, J., “Methodological framework for data processing based on the Data Science paradigm,” in Proceedings of the 2014 Latin American Computing Conference, CLEI 2014, 2014. https://doi.org/10.1109/CLEI.2014.6965184

[21] Chembessi, C., Beaurain, C., and Cloutier, G., “Analyzing technical and organizational changes in Circular Economy (CE) implementation with a TOE framework: insights from a CE project of Kamouraska (Quebec).”, Circ.Econ.Sust. 2022, Vol. 2, 915–936, https://doi.org/10.1007/s43615-021-00140-y

[22] Samarasinghe, S., and Lokuge, S., “Data-driven Innovation: Understanding the Direction for Future Research.”, ACIS 2022, 2022, Proceedings. 16, https://doi.org/10.48550/arXiv.2212.03061

[23] Babu, M. M., Rahman, M., Alam, A., and Dey, B. L. “Exploring big data-driven innovation in the manufacturing sector: evidence from UK firms,” Annals of Operations Research, 2021, https://doi.org/10.1007/s10479-021-04077-1

[24] Wang, X., Liu, Y., Zhang, H., Ma, Q., & Cao, Z. “Public Health Emergency Management and Multi-Source Data Technology in China”. Intelligent Automation and Soft Computing, 2018, Vol. 24, no. 1, pp. 89–96. https://doi.org/10.1080/10798587.2016.1267246

[25] Koukaras, P., Nousi, C., Tjortjis, C., “Stock Market Prediction Using Microblogging Sentiment Analysis and Machine Learning”. Telecom, 2022, Vol. 3 no. 2, pp 358-378, https://doi.org/10.3390/telecom3020019

[26] Aguilar J., Buendia O., Pinto A., Gutiérrez J. “Social learning analytics for determining learning styles in a smart classroom”, Interactive Learning Environments, 2022, Vol. 30, no. 2, pp. 245-261, https://doi.org/10.1080/10494820.2019.1651745

[27] Sánchez M., Aguilar J., Cordero J., Valdiviezo-Díaz P., Barba-Guamán L., Chamba-Eras L. “Cloud Computing in Smart Educational Environments: Application in Learning Analytics as Service”. In: New Advances in Information Systems and Technologies (Rocha, Á., Correia, A., Adeli, H., Reis, L., Mendonça Teixeira, M. (eds)), 2016, Vol 444. pp 993–1002. https://doi.org/10.1007/978-3-319-31232-3_94

[28] Aguilar J., Garcès-Jimènez A., Gallego-Salvador N., De Mesa J., Gomez-Pulido J., Garcìa-Tejedor À. "Autonomic Management Architecture for Multi-HVAC Systems in Smart Buildings," IEEE Access, 2019, Vol. 7, pp. 123402-123415, https://10.1109/ACCESS.2019.2937639.

[29] Morales, L., Ouedraogo, C.A., Aguilar, J. *et al.* “Experimental comparison of the diagnostic capabilities of classification and clustering algorithms for the QoS management in an autonomic IoT platform”. *Service Oriented Computing and Applications* 2019, Vol. 13, pp. 199–219.

[30] Koukaras, P., Nousi, C., Tjortjis, C., “Stock Market Prediction Using Microblogging Sentiment Analysis and Machine Learning”. Telecom, 2022, Vol. 3 no. 2, pp 358-378, https://doi.org/10.3390/telecom3020019

[31] Fornieles, R., “De Lasswell a Gorgias: los orígenes de un paradigma”, Estudios sobre el Mensaje Periodístico, 2012, Vol. 18, no. 2, https://doi.org/10.5209/rev_esmp.2012.v18.n2.41043

[32] Singh L., and Singh, S., “Empirical study of sentiment analysis tools and techniques on societal topics,” Journal of Intelligent Information Systems, 2021, Vol. 56, no. 2, pp. 379–407, https://doi.org/10.1007/s10844-020-00616-7

[33] Oad, A., Koondhar, I., Butt, P., Lei, H., Oad, A., Khuhro, M., and Bhutto, S., “VADER Sentiment Analysis without and with English Punctuation Marks,” International Journal of Advanced Trends in Computer Science and Engineering, 2021, Vol. 10, no. 2, pp. 1483–1488, https://doi.org/10.30534/ijatcse/2021/1371022021

[34] Sánchez, J., “Humor, ironía y sarcasmo en las redes sociales: creación léxica a partir del formante grecolatino -itis”, Revista de Investigación Lingüística, Vol. 25, pp. 259–283. https://doi.org/10.6018/ril.482871

[35] Oliveira, G., Coutinho, F., Campello, R., and Naldi, M. “Improving k-means through distributed scalable metaheuristics,” Neurocomputing, 2017, Vol. 246, pp. 45–57, https://doi.org/10.1016/j.neucom.2016.07.07

[36] Hong, T., Lin, C., Yang, K., and Wang, S. “Using TF-IDF to hide sensitive itemsets”. Appl Intell. 2013, Vol. 38. pp. 502-510. https://doi.org/10.1007/s10489-012-0377-5.

Appendix E

Using Fuzzy Cognitive Maps to evaluate the innovation in Micro, Small and Medium Enterprises

Using Fuzzy Cognitive Maps to evaluate the innovation in Micro, Small and Medium Enterprises

Ana Gutiérrez[1,2,5], Jose Aguilar[2,3,4,5,*], Ana Ortega[2], Edwin Montoya[2,5]

[1] GIA, Universidad Francisco de Paula Santander, Av. Gran Colombia, Cúcuta, Colombia

[2] GIDITICS, Universidad EAFIT, Medellín, Colombia

[3] CEMISID, Universidad de Los Andes, Núcleo La Hechicera, Mérida, Venezuela

[4] IMDEA Networks Institute, Leganés, Madrid, Spain

[5] Tepuy R+D Group. Artificial Intelligence Software Development. Mérida, Venezuela

[*]Corresponding author

Abstract. The field of innovation covers not only the development of new products, but also new internal processes, the management of new market positioning initiatives, and the generation of new concepts and platforms, sometimes based on new paradigms. In the innovation process, organizations should measure activities and evaluate innovation performance. This will allow them to continuously improve their processes, methods, and capabilities, and increase product and process innovation in the future. Controlling the evaluation of innovation using intelligent systems can be challenging for companies seeking innovation performance. This paper presents a fuzzy cognitive map for improving the assessment of innovation in Micro, Small and Medium Enterprises. The previous results in the literature show that although there are relevant advances on this topic, there is not enough knowledge to provide clear guidelines for evaluating innovation and improving performance in an organization using intelligent systems. However, we get very good results with our approach based on fuzzy cognitive maps in three case studies to determine the level of innovation in an organization. The fuzzy cognitive maps in model demonstrated a high level of accuracy, with an 80% accuracy rate for case studies from Colombia and an impressive 92% accuracy rate for global case studies. This indicates the model's effectiveness in evaluating innovation levels in organizations quantitatively. Furthermore, the study found that innovative activities emerged as an influential factor in organizations, essential to improve their operations and competitiveness. Additionally, this article reveals and addresses several challenges in this field.

Keywords: Innovation Assessment, Artificial Intelligence, Intelligent Systems, Micro-Small and Medium Enterprises, Machine Learning, Fuzzy Cognitive Maps

1. Introduction

Innovation is a phenomenon that has been studied from different perspectives, bearing in mind that it is considered as a process but also as a result; and as a process, it encompasses different elements and stages of development. The interest in the subject from the theoretical, academic and practical perspectives derives from the importance of innovation for economic and social development at different levels (e.g., national, regional, organizational), as well as for its role in the solution of the current social problems. Considering its importance, various strategies have been developed to measure and quantify innovation activities and results. Evidence of this is the Oslo Manual [1] [2], which aims to provide guidance on the collection and interpretation of innovation data.

An adequate innovation evaluation process allows assessing the innovation capacity within the organizations, in such a way that both strengths and weaknesses in this field can be identified, enabling the incorporation of improvements and the implementation of orderly and systematized plans to achieve the innovation objectives, promoting and driving innovation in organizations.

The purpose of this paper is to develop a tool to evaluate the innovative capacity of organizations based on Fuzzy Cognitive Maps (FCM), particularly, for improving the evaluation of innovation in Micro, Small and Medium Enterprises (MSME). Particularly, a tool like this within a firm allows identifies future actions (processes and activities) that will help the organization to achieve the expected innovation performance. This work differs from previous research since this article focuses on how intelligent systems have been used to improve the evaluation of organizational innovation, understanding innovation evaluation as a different way of improving innovation for MSME, an approach that has not been used before.

1. Related works

In general, the FCMs have been used in numerous domains [3], [4]. This section presents the most important recent research related to the use of FCMs in processes related to innovation assessment in organizations.

In the article [7], Wie introduces FCMs for the modeling and evaluating of trust relationships, and explores the sources of trust and the relationships between organizations based on their veracity. There are two main contributions of this research. First, the addressing of the critical trust factors in the virtual enterprise (VE) context considers an inter-organizational environment, where team trust is a function of three key elements: trust giver, trustee, and environment. Second, this paper proposes a methodology that considers the dynamic nature of trust to analyze the evolution of trust in VE environments [7]. Thus, this study shows how to apply FCM for the assessment in a VE.

Also, in our attempt to apply FCM to innovation, we found useful the work of Kok [8], who presented the potential improvements to FCM methodology [8] to be used in an organization, by presenting a fuzzy psychological function map, critically assessing its added value in the context of situational development. Aguilar [5] presented an FCM, which was built on the principles of a random neural network, it was called a *random cognitive fuzzy map,* a technique that showed its application in different modeling processes in an organization [6]. There are other interesting works on FCM, extensions of the model, new learning algorithms, new models of causality, etc. [9][10][11]. For example, Aguilar [12] suggests a related extension of FCM, allowing connections between multiple FCMs. This extension applies to the FCM designer tool to facilitate future FCM development and execution. For the present work, this tool was used to develop our FCM in relation to innovation assessment.

Only fourteen previous works in the literature used an FCM tool for executing and simulating a real situation with real data. The main areas where the previous works have been applied are in business [7][13][14][15], in an environment [8], in the industry [6], in the computer sciences [5][10][11][12][16], in medicine [17][18], and social science [15] (see a summary in Figure 1). The previous works on FCM show the key aspects that we need to consider in our work. Due to its flexibility in the application, it becomes a suitable approach for the investigation, allowing the use of the FCM for the innovation level assessment.

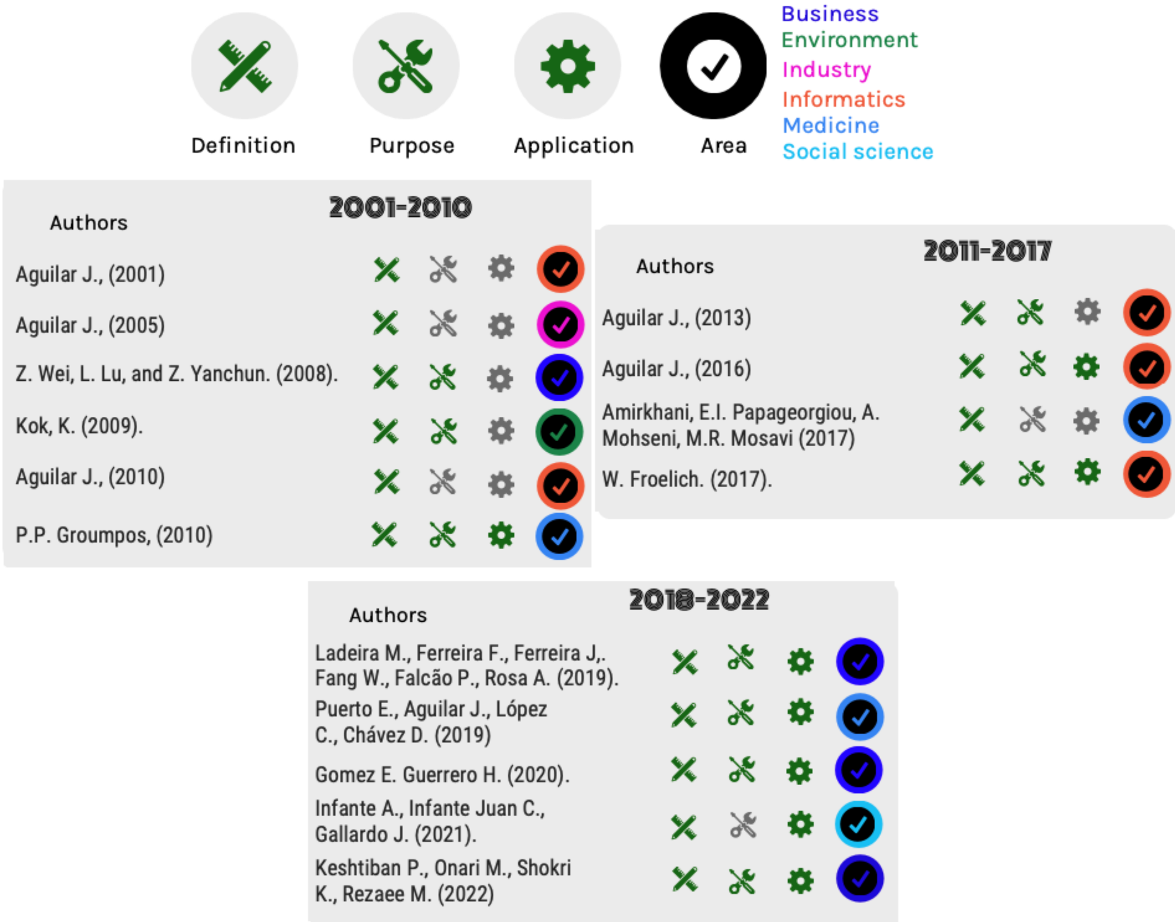

Figure 1. Literature Review Summary - FCM

Previous works have not defined a framework for evaluating innovation in MSMEs based on FCMs. They also do not use the data of an organization to assess the key characteristics related to innovation. This work applies FCM to automate the evaluation of the process and the capacity for innovation in an organization. These are the main differences between our approach from previous studies. The main motivation behind choosing FCMs lies in their capacity to model complex, dynamic, and uncertain relationships within the innovation process effectively. Their cognitive modeling approach, adaptability, and diagnostic capabilities make them a valuable tool for understanding, evaluating, and improving innovation management within organizations. These combined properties make this system more robust.

1.2 Contribution of this work

The contribution of this work consists of the design of an FCM that enables the evaluation of innovation, and to achieve this, the main aspects for the design of FCMs will be applied, especially, by identifying the concepts and their relationships. We use real data from different sources in the world, derived from real situations, to demonstrate the quality of our FCM. The high accuracy rates achieved by the FCM model (80% for Colombian case studies and 92% for global case studies) suggest that organizations can rely on this tool to quantitatively assess their innovation levels more precisely. This tool can help in identifying areas that require improvement to impact positively organizations. By improving innovation assessment through the FCM model, organizations can anticipate higher profitability because innovations are often closely tied to revenue generation and cost savings. The tool can determine the necessity of new products or services, improve operational processes, or enter new markets. Finally, the paper analyzes current challenges related to the application of FCM to improve MSME innovation assessment.

This work is organized as follows: the second section presents the theoretical framework for innovation assessment and FCM and its application in the innovation domain. Also, this section reviews the most important recent research related to the use of FCMs in processes related to innovation assessment in organizations. Then, section 3 describes the design of the FCM. The fourth section describes the main results of the experiments. The fifth section presents comparisons with other works. Finally, in section 6, conclusions are presented, outlining the contributions of this work, and some recommendations for future studies.

2. Theoretical background

This section presents a background about innovation assessment and the main basic concepts of the FCMs and their utilization in the innovation domain.

2.1 Innovation

It is difficult to think about moving forward and staying in the market without innovation now. Regardless of the type of company, the success and development of the organization can be attributed to its innovative activities and results [19]. SMEs are the sectors of business life that need the most constant innovation to improve their business [20]. Thus, innovation is generally seen as a key factor in the survival and success of a company as it needs to adapt more quickly and better to the needs of a changing environment. The importance of innovation is often made more evident by rapid change, which requires managers to research and implement activities that enable them to respond to market needs and design and implement appropriate strategies in their companies [22].

On the other hand, SMEs often avoid implementing a comprehensive value plan in terms of innovation, technology management, and R&D investment [21]. In this way, the organizational leaders of this type of organization must understand the importance of innovation and address the necessary internal changes related to human capital development and creativity [22].

2.1.1 Oslo Manual

The Oslo Handbook is the foundation upon which the OECD and other international organizations collect and publish statistics on business innovation. The latest is his 2018 edition, which includes improved guidelines that reflect changing user interests and accumulated real-world experience. It includes new materials to support the measurement of innovation outside the firm, understanding the internal and external drivers of innovation in the firm and the key innovations in the business, and making better use of innovation data in statistics and analytics. It is an international benchmark for multiple works in the innovation domain [1].

The Oslo Manual defines a set of concepts to measure the level of innovation in an organization (see Table 1) [1]. These concepts are grouped into four categories: i) innovation types, ii) innovation activities, iii) basic research, iv) external context.

Table 1. Concepts for innovation evaluation according to the Oslo handbook [1].

Category	**Concept name**	**Description**
Innovation Types	Production	Engineering activities and related technical testing, analysis and certification to support production
	Distribution	It includes new marketing methods or new organizational methods.
	Engineering and	Engineering involves procedures, methods and standards of production and quality

	design	control
	Product	The organization has strategies to create, improve or update its products, goods or services.
	Distribution and logistics	Transportation, storage and order processing
	Information and communication systems	Hardware and software sales and logistics. data processing and databases; maintenance; web hosting and other activities.
	Administration and management	Includes: strategic business management, corporate governance, accounting and purchasing
	Process	The organization has agile strategies to generate significant changes in its processes
Innovative Activities	Customers / Users	The organization receives feedback and suggestions from customers regarding satisfaction with the use of the products offered by the company
	Competence	The organization has the competence to carry out organizational surveillance tasks to feed the innovation processes in the organization
	Standards	The organization considers quality standards and manuals and guides for innovation processes when implementing them within the organization.
	New Products	The organization has activities to develop new products, functions, or formats of products and processes
	Marketing	The organization has the flexibility to implement new marketing methods to reach more customers, changes in design, packaging, and promotion
	External Knowledge	Information sources, knowledge, technology, best practices, and human and financial resources
	Agents of the innovation system	Official laboratories, Universities, ministerial departments, regulatory authorities, competitors, suppliers, and customers.
	Promotion	The organization carries out studies and tests on the market, promotion, and start-up of the products
	Employee training	Job training and job-related education in educational and training institutions

	Staff	The organization has trained staff that encourages innovation processes
	Secondary education	Net coverage rate in middle and secondary education. School life expectancy and Evaluation of school performance in reading, mathematics, and science in the Saber 11 tests
	Knowledge	The organization has personnel who know how to start, develop, and achieve innovative results
	Experience	The staff of the organization has experience in innovation processes, implementation of innovation proposals
	Techniques	The organization has personnel who know the techniques of Innovation Design Thinking, co-creation, rapid prototyping, Personal
	Innovation Management	Plan, govern and control internal and external resources for innovation
	Training	The organization carries out staff training activities to use innovations, or to implement innovative software
	Digital	The organization has personnel trained in Business Intelligence, Data Analytics, to improve or develop the business model.
	Use	The organization has staff who have skills to incorporate and use new technology
	Intellectual Property	The organization's staff has the skills to manage the intellectual property of innovation, Technological tools
	Design	The organization has personnel who are capable of designing, making technical specifications, prototype
Basic Research	Research and experimental development	It includes systematic reviews and creative work done to expand knowledge, develop new applications in order to do available knowledge
	Human Resources External	The organization has external personnel available trained to support the Innovation process (consultants, R&D companies, public S&T agencies, Chambers of Commerce)
	Size of the company	Firm size is a commonly used predictor of innovation activities and firm

	Acquisition or Lease of Property, Plant and Equipment	Innovation activities exist when these assets are necessary for product or business process innovation.
	Investment	Innovation implies including acquisitions of tangible and intangible assets as well as any other type that can influence the efficiency of your innovation activities
	Protection	The organization executes activities to protect ideas, improved products, or new processes against copying by its competitors
	Development	The organization develops and uses software to implement new or improved products or process
	Assets	The organization has a physical plant and assets that are used in the development of innovation processes.
	Source of funds	Own funds, transfers, customer orders, shareholder loans, debt financing, government loans, international organizations, among others
	Partners	The partners of the organization support and support the innovation processes
	Finance	The organization has sources of finance available to promote innovation processes Departures
External context	Diffusion	Whereby innovations spread through commercial circuits, or any other, to different consumers, countries, regions, sectors, markets, and companies
	Networks	The organization has alliances with other companies in the productive sector that promote the innovation process
	Digital Platforms	Is the organization active on digital platforms? For example, it has a website, it has a presence on social networks, Market Tickets – Sources
	Location	The organization is strategically located in an environment of innovation
	Public Infrastructure	The organization's environment has public infrastructure and road networks that allow it to develop innovation processes
	Macroeconomic Policy	The organization considers and applies legislation on patents, taxes, regulations on company management
	Taxes	The environment where the organization is located has fiscal and tax policies that encourage innovation processes.

	Environmental sustainability	Efficiency in the use of energy (in units of GDP), Environmental performance (0-100), Companies with environmental certification ISO 14001
	Government support	The government where the organization is located has policies, strategies that promote innovation

2.2 Fundamentals of FCM

FCMs are methods for modeling systems based on the causal relationships between their concepts, which can be positive or negative. FCMs were proposed by Kosko in 1986 [3]. FCMs are derived from the cognitive maps [23] created by Axelrod in 1976, and are used to model complex systems in terms of ease of construction and interpretation. They are also called directed graphs, and consist of concepts and relationships between them (se Figure 1). These concepts and their relationships represent fuzzy variables that can be expressed in the linguistic terms perfect, high, medium, and low [11].

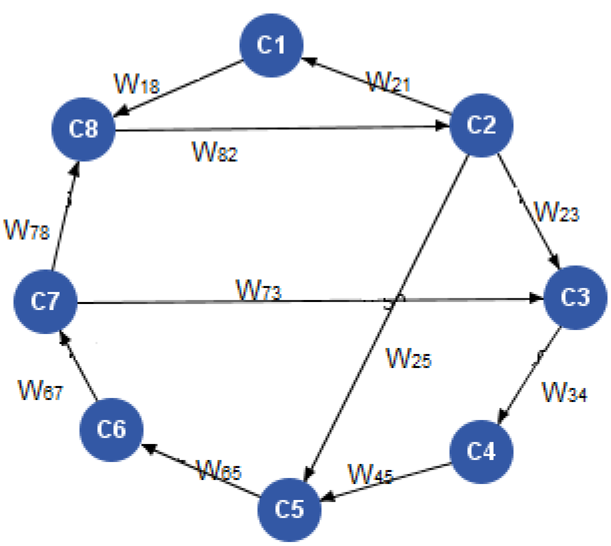

Figure 2. An FCM with eight concepts

Fig. 2 represents a simple FCM with eight concepts (from C_1 to C_8). Each concept constitutes a variable or element of the system. A relationship between them is represented by a directed edge with a weight, to define the causal relationship between them [6]. The value of a relationship can vary from -1 to +1. The value +1 means an excitatory effect from a concept C_i to a concept C_j. The value -1 represents an inhibited effect between C_i and C_j. The value 0 defines that there is no causal relationship between the two concepts. Finally, other values between (-1 and +1) symbolize different degrees of causality. An FCM is formally defined as a tuple:

$$\Phi = \langle n, W, f\,(.)\rangle \tag{1}$$

Where $\boldsymbol{n} \in \boldsymbol{R^m}$ is the set of concepts $(n_1, \ldots, n_m)$; $W \epsilon R^{mxm}$ is the adjacency matrix that will symbolize the interactions between concepts; f (.) is the activation function that defines the concept values, which can be in the range [0, 1] for a sigmoid function or [-1, +1] for a hyperbolic tangent function [24]. An adjacency matrix is a square matrix that defines the causal relationships between the concepts of an FCM. As an example, the adjacency matrix for the FCM in Fig. 1 is:

$$\mathbf{W} = \begin{array}{c} \\ \boldsymbol{C_1} \\ \boldsymbol{C_2} \\ \boldsymbol{C_3} \\ \boldsymbol{C_4} \\ \boldsymbol{C_5} \\ \boldsymbol{C_6} \\ \boldsymbol{C_7} \\ \boldsymbol{C_8} \end{array} \begin{array}{c} \begin{array}{cccccccc} \boldsymbol{C_1} & \boldsymbol{C_2} & \boldsymbol{C_3} & \boldsymbol{C_4} & \boldsymbol{C_5} & \boldsymbol{C_6} & \boldsymbol{C_7} & \boldsymbol{C_8} \end{array} \\ \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & w_{18} \\ w_{21} & 0 & w_{23} & 0 & w_{25} & 0 & 0 & 0 \\ 0 & 0 & 0 & w_{34} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & w_{54} & 0 & w_{56} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & w_{67} & 0 \\ 0 & 0 & w_{73} & 0 & 0 & 0 & 0 & 0 \\ 0 & w_{82} & 0 & 0 & 0 & 0 & w_{87} & 0 \end{pmatrix} \end{array}$$

The inference process in an FCM can be specified from the mathematical point of view from two components: a state vector, $\boldsymbol{a} \in \boldsymbol{R^m}$, which means the degree of activation of the concepts, and a weight matrix, $\boldsymbol{W} \in \boldsymbol{R^{mxm}}$, which specifies the interaction between the concepts. Thus, the inference process consists of calculating the state vector in a given time, from an initial condition a(0). The activation vector $a(0)$ is:

$$a(0) = \left(a_1(0), a_2(0), \dots, a_m(0)\right) \quad (2)$$

Where a_1 (0) is the value of the concept C_1 at time 0.

The activation of C_j will depend on the activation of the concepts that directly affect C_j and the causal weights associated with that concept. The following equation summarizes this process:

$$a_j(t+1) = f\left(\sum_{i=1, i\to j}^{m} W_{ij} a_i(t)\right) \quad (3)$$

Where $a_j(t+1)$ is the value of concept C_j at time t+1. Normally, an inference process iterates until the system finds an equilibrium point. This equilibrium point (final state) is reached when a(t) = a(t-1), or a(t)-a(t-1)≤0.001.

3. Design of the innovation assessment system using FCMs

This section presents the specification of the innovation assessment system based on FCMs. For that, this work uses the methodology called Intelligent Decision Support System (IDSS) [27]. Fig. 3 shows the six-step methodological framework. Each step is briefly described below.

3.1. Selection of experts

Eight people participated, as they were considered experts in innovation within the organizations, who agreed to participate in this study. The experts were selected because they have extensive experience in the Oslo manual and innovation assessment processes.

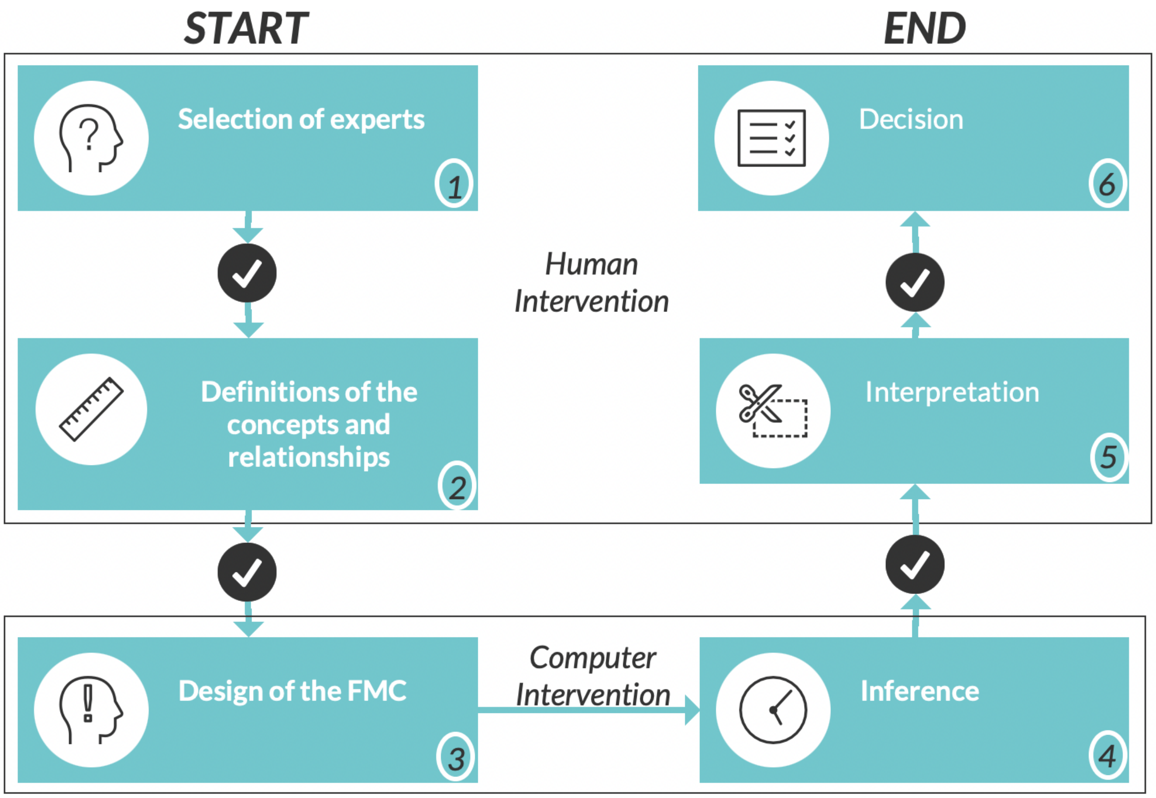

Fig. 3. Flowchart of the IDSS for the development of the innovation assessment system.

3.2. Concepts and relationships

The concept was defined according to the Oslo Manual [1] (see Table 1), such that each concept of the Oslo Manual is a concept of the innovation assessment system. Table 2 shows the concept used by our FCM.

Table 2. Concepts used by the innovation assessment system.

Node	Concept name
C1	Production
C2	Distribution
C3	Engineering and design

C4	Product
C5	Distribution and logistics
C6	Information and communication systems
C7	Administration and management
C8	Process
C9	Customers / Users
C10	Competence
C11	Standards
C12	New Products
C13	Marketing
C14	External Knowledge
C15	Agents of the innovation system
C16	Promotion
C17	Employee training
C18	Staff
C19	Secondary education
C20	Knowledge
C21	Experience
C22	Techniques
C23	Innovation Management

C24	Training
C25	Digital
C26	Use
C27	Intellectual Property
C28	Design
C29	Research and experimental development
C30	Human Resources External
C31	Size of the company
C32	Acquisition or Lease of Property, Plant and Equipment
C33	Investment
C34	Protection
C35	Development
C36	Assets
C37	Source of funds
C38	Partners
C39	Finance
C40	Diffusion
C41	Networks
C42	Digital Platforms
C43	Location

C44	Public Infrastructure
C45	Macroeconomic Policy
C46	Taxes
C47	Environmental sustainability
C48	Government support

After defining the concept, each expert used their knowledge and experience to create a weight matrix (values between -1 and 1) corresponding to the impact between the concepts of the FCM. Possible types of relationships were high, medium, and low, depending on the causality between the concepts.

3.3. Design of the model

This stage summarizes the FCM developed from the different opinions of the experts, in order to generate a general FCM. The procedure to create a single FC< is defined by the following equation [5]:

$$E_{ij}^{G} = \sum_{e=1}^{NE} \frac{E_{ij}^{e}}{NE} \qquad (4)$$

where $E_{ij}^{G}$is the global weight of the general FCM, E_{ij}^{G} is the opinion of each expert on the

causal relationship between the concept C_i and C_j, and NE is the number of experts. Finally, the general FCM is built using the FCM Expert, which was developed by Nápoles in 2017 [10], which is a software platform for modeling FCM based systems. Figure 4 shows the general FCM.

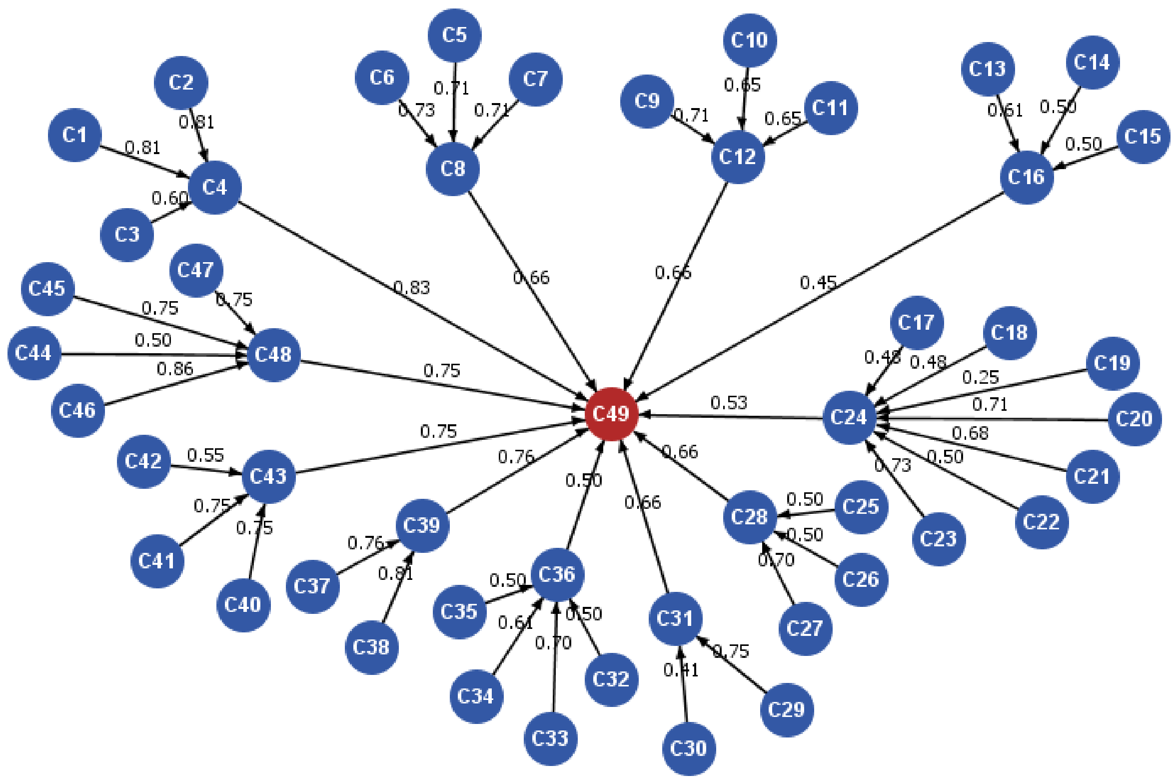

Fig. 4. FCM of the Innovation Assessment System.

Thus, step 3 of the methodology, shown in Figure 5, consisted of implementing the previously defined FCM in the FCM tool [10] used in this work. Step 4 is the inference process for each case study. Finally, steps 5 and 6 interpret the inferred results by the FCM regarding innovation ratings for each case study and analyze the causal relationships used by the FCM to derive these ratings. The next section describes these steps.

4. Experiments

In this section, several case studies were studied. In the first case study is used data about the Departmental Innovation Index for Colombia. The second case study uses the World Innovation Index. For both case studies, the data for the last five years (from 2018 until 2022). The last case study uses data from the MSMSs based in the Valle de Aburrá, specifically in Ruta N [28].

On the other hand, to evaluate the quality of the FCM to determine the level of innovation, two metrics were used: the ability to classify the known data well and the Cohen's Kappa Coefficient. Cohen's Kappa Coefficient measures the agreement between two raters who each classify N items into C mutually exclusive categories. If the raters are in complete agreement, then kappa =1. If there is no agreement among the raters, then kappa =0.

4.1. Departmental Innovation Index for Colombia Case Study

This dataset corresponds to the registry of innovation of each one of the departments of Colombia during the years 2018 to 2021 [25]. The variables in the dataset are associated with types of innovation, innovative activities, basic research, and external context. For this case, each of the values of the variables considered in the case study was associated with a concept in the FCM (see Table 3).

Table 3. Relationship among concepts of the innovation assessment system and variables of the Departmental Innovation Index dataset.

Node	Concept in the FCM	Variable in the Dataset [25]
C2	Distribution	Intangible goods
C3	Engineering and design	Knowledge impact
C15	Agents of the innovation system	Innovation links
C16	Promotion	Trade and competition
C17	Employee training	Higher education
C19	Secondary education	Secondary and middle education
C20	Knowledge	Knowledge workers
C21	Experience	Knowledge creation
C24	Training	Online creativity

C29	Research and experimental development	Investigation and development
C33	Investment	Knowledge absorption
C34	Protection	Creative goods and services
C36	Assets	Available resources
C39	Finance	Credit
C40	Diffusion	Knowledge diffusion
C42	Digital Platforms	TIC
C44	Public Infrastructure	General infrastructure
C45	Macroeconomic Policy	Regulatory environment
C46	Taxes	Business environment
C47	Environmental sustainability	Environmental sustainability
C48	Government support	Political environment

As a result of the previous process, a new FCM with 22 concepts is obtained. Finally, the number of data in this dataset is 2621.

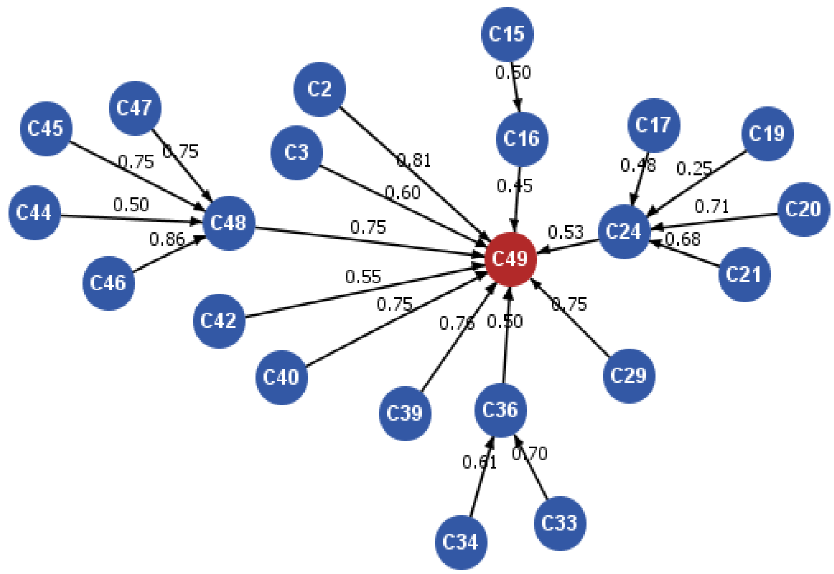

Fig. 5. FCM for the dataset of the Departmental Innovation Index for Colombia.

In all the tests of this case study, 80% of the data was left for training, and the rest was to evaluate the capacity of the FCM to determine the level of innovation in each organization. In the first test of this case study, the level of innovation was defined by 5 classes: Low Level, Low Medium Level, Medium Level, High Medium Level, and High Level. The prediction results of the level of innovation in each organization are shown in Table 4.

Table 4. Prediction results for the first test in the first case study.

Criterion	**Results**
Correctly Classified Instances	52.0%
Incorrectly Classified Instances	48.0%
Cohen's Kappa Coefficient	0.38

The correct prediction of the FCM is only 52%. According to the results, classes 0 and 1 were the ones that obtained a higher percentage of correct answers in the classification (see Table 7). For classes 2, 3 and 4, the results for the classification were always wrong.

In a second test, the values for the classification of the level of innovation are redefined, leaving only three classes: Low, Medium, and High. The results of this second test for the first case study are shown in Table 5.

Table 5. Prediction results for the second test in the first case study.

Criterion	**Results**
Correctly Classified Instances	60.0%
Incorrectly Classified Instances	40.0%
Cohen's Kappa Coefficient	0.41

In the third test, we balance the classes. In this case, the results improved considerably with respect to those obtained before (see Table 6).

Table 6. Prediction results for the third test in the first case study.

Criterion	**Results**
Correctly Classified Instances	82.0%
Incorrectly Classified Instances	18.0%
Cohen's Kappa Coefficient	0.68

In this case, 80% of the data are correctly classified. The following table shows the precision results for each class in the three tests.

Table 7 Accuracy results of the tests carried out by class.

Class	**Test 1**	**Test 2**	**Test 3**
0	0.41	0.68	0.91
1	0.75	0,65	0.81

2	0	1	0.7
3	0	-	-
4	0	-	-

According to the results (See Table 7), it is evident that the accuracy is improved as the classes are redefined and balanced. In the first test, with five classes, the FCM failed to correctly predict classes two, three and four. For tests two and three, the accuracy improves, but in particular, it is much better with balanced classes (case 3).

4.2. Global Innovation Index Case Study

In this second case study, the dataset corresponds to the registry of the countries of the world during the years 2018 to 2022 [29]. The variables in this dataset are associated with the concepts of Institutions, Human capital and research, Market sophistication, Business sophistication, Knowledge and technology and creative outputs. For this case, each of the variables of the dataset was associated with a concept in our FCM (see Table 8).

Table 8. Relationship among concepts of the innovation assessment system and variables of the Global Innovation Index dataset.

Node	**Concept in the FCM**	**Variable in the Dataset**
C2	Distribution	Creative production
C3	Engineering and design	Production of knowledge and technology
C15	Agents of the innovation system	Business sophistication
C16	Promotion	Market sophistication
C17	Employee training	Human Capital and Research

C19	Secondary education	Human Capital and Research
C20	Knowledge	Business sophistication
C21	Experience	Production of knowledge and technology
C24	Training	Creative production
C29	Research and experimental development	Human Capital and Research
C33	Investment	Market sophistication
C34	Protection	Creative production
C36	Assets	Business sophistication
C39	Finance	Market sophistication
C40	Diffusion	Production of knowledge and technology
C42	Digital Platforms	Infrastructure
C44	Public Infrastructure	Infrastructure
C45	Macroeconomic Policy	Institutions
C46	Taxes	Institutions
C47	Environmental sustainability	Infrastructure
C48	Government support	Institutions

As a result of the previous process, a new map is defined with seven main concepts, each encompassing several concepts. These main concepts are Institutions (C1), Human Capital and Research (C2), Infrastructure (C3), Market Sophistication (C4), Business sophistication (C5), Production of knowledge and technology (C7), and Creative Production (C8). And for example, the main concept of Institutions (C1) corresponds to the following three concepts of our original model: Macroeconomic Policy, Taxes, and Government support. Finally, the number of data was 3626.

With this redefinition of the FCM and the data, 80% of them are used to train and thus define the relationships between the concepts. The resulting FCM using the historical data from [29] is the one presented in Figure 6. The FCM also shows that C8 corresponds to the decision node, which is the level of innovation predicted (this concept is classified in three classes: Low, Medium, and High).

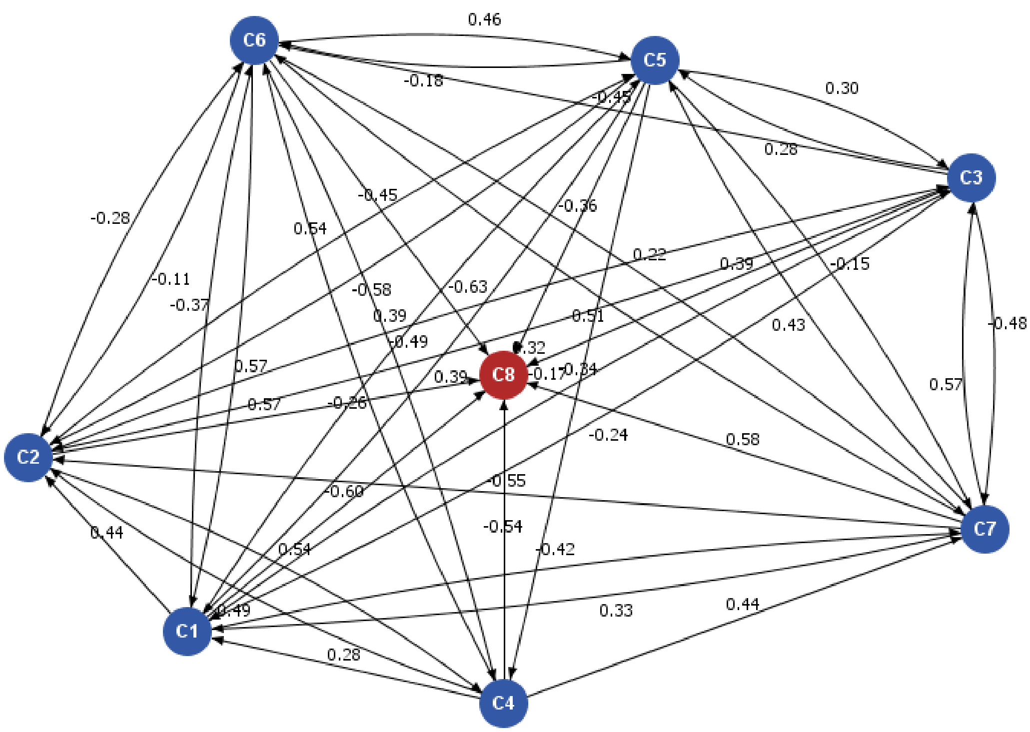

Fig. 6. FCM of the Innovation Assessment System for the Global Innovation Index dataset.

Table 8 shows the results, which are very good. The correct prediction of the FCM (more than 92%) with a very good value of Cohen's Kappa Coefficient.

Table 8. Prediction results for the second case study.

Criterion	**Results**

Correctly Classified Instances	92,2
Incorrectly Classified Instances	7,7%
Cohen's Kappa Coefficient	0.88

4.3. Route N Case Study

In the latter case, we work with the information provided by the "Ruta N" Corporation in Colombia, which is the entity in charge of registering and monitoring the efforts and achievements of MSMEs in Information Technology located in the Aburrá Valley, Colombia. This corporation provided us with a dataset on the innovation processes in these organizations, based on surveys that they carry out periodically to MSMEs [28]. For the creation of the FCM, the concepts related to each of the survey questions were identified, which is shown in Table 9.

Table 9. Relationship among concepts of the innovation assessment system and variables of the survey of Route N.

Node	**Concept in the FCM**	**Variable in the Dataset**
C4	Product	New goods or enhanced to offer market, methods or techniques news for your processes
C10	Competence	Form distribution of Organization to lead innovation processes
C12	New Products	Number of new Activities, goods or services in the national market
C15	Agents of the innovation system	Use of innovation links
C17	Employee training	Academic degrees (or are currently studying)
C19	Secondary education	Academic levels or grades
C20	Knowledge	People with responsibilities for innovation activities

C21	Experience	Professionals, Researchers or technicians who work in Research and Development
C23	Innovation Management	Management supports and promotes innovation
C27	Intellectual Property	The organization has in process of any of the intellectual property registrations
C34	Protection	Protect these products/services/methods, is it did you use any mechanism?
C36	Assets	Total assets
C37	Source of funds	Capital to finance innovation activities
C39	Finance	Funds for the development of innovation activities
C41	Networks	Relationships with other entities

The FCM is visualized in Figure 7 after the training phase using 80% of the data from the dataset [28]. Previously, the variables were normalized in a range between (0) and (1) for the training process.

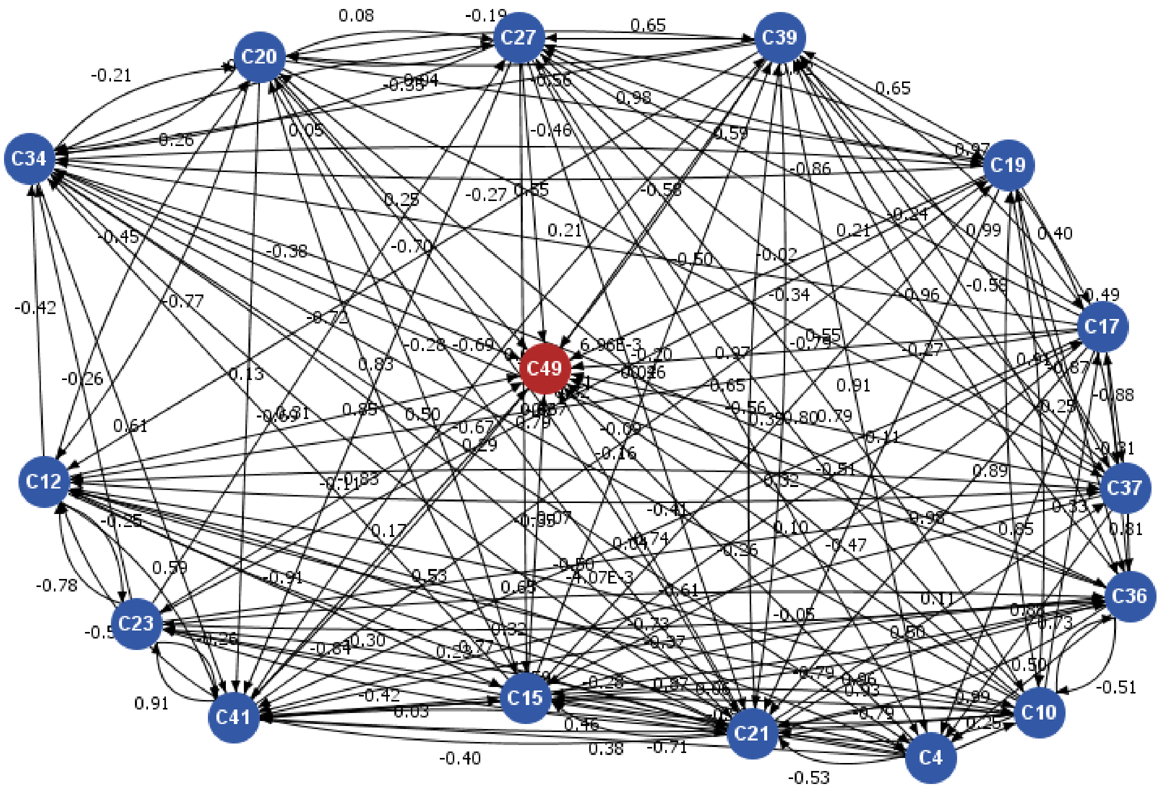

Figure 7. FCM for Route N case study

Figure 7 shows the fifteen concepts that correspond to the variables of the dataset. FCM also shows C49 that is the decision node, which is the level of innovation predicted, defined in three classes: Low, Medium, and High.

Table 10 shows the results of this case study, which are good. The correct prediction of the FCM is (77%), with a good Cohen's Kappa Coefficient.

Table 10. Prediction results for Route N case.

Criterion	Results
Correctly Classified Instances	77,0%
Incorrectly Classified Instances	23,0%
Cohen's Kappa Coefficient	0.68

5. Comparison with previous works

In this section, a comparison is made to evaluate our proposal based on FCM in the evaluation of the level of innovation in organizations with different previous works. For this purpose, the following criteria are described:

Criterion 1: Based on the Oslo Manual standard.

Criterion 2: Uses automated methods.

Criterion 3: Uses artificial intelligence.

Criterion 4: Evaluates innovation.

Criterion 5: Uses FCMs.

Table 11 provides a qualitative assessment with related work based on these criteria.

Table 11. Comparison with previous work

Work	Criterion 1	Criterion 2	Criterion 3	Criterion 4	Criterion 5
[7]	✗	✓	✓	✓	✓
[25]	✗	✗	✗	✓	✗
[26]	✗	✗	✓	✓	✗
This work	✓	✓	✓	✓	✓

.

According to criterion 1, the other works do not consider the Oslo Manual standard, which is only used in our research. This manual "is part of a series of evolutionary manuals devoted to the measurement and interpretation of data related to science and technology in innovation" [1], hence this is an originality of our research in comparison with previous studies.

Regarding criterion 2, it is observed that Wei [7] also uses the FCMs for the evaluation of innovation, but in the present work, the conceptualization is carried out using the Oslo manual [1]. According to the results, our FCM leads to a classification whose margin of error is small. Therefore, FCMs are very useful for the evaluation of the innovation in companies, which cannot be seen in the research of Wei [7].

Related to criterion 3, in addition to Wei [7], other works use artificial intelligence, such as [26] which proposed an evaluation through a process based on statistical techniques of artificial intelligence in various industrial sectors of Europe and America. On the other hand, when referring to criterion 4, all the authors described above applied this criterion for evaluating innovations. However, other authors follow the traditional methodology to evaluate innovations in organizations.

Finally, regarding criterion 5, FCMs were only used by Wei [7] and in our research. It is noteworthy that the evaluation of innovations in companies has been developing for years, and the methods used have been evolving. Thus, our work represents a great contribution, and a great novelty is the evaluation of innovations based on the Oslo Manual.

Finally, the utilization of FCMs in innovation assessment promises to significantly improve the management of innovation within organizations. By capturing complex relationships among innovation-related factors, it enables more strategic and informed decision-making. The FCM model's high accuracy rates (80% for Colombian cases and 92% for global cases) suggest that organizations can rely on this tool to quantitatively assess innovation levels more precisely. This, in turn, helps organizations understand where they stand in terms of innovation and what steps they need to take for improvement.

By strategically allocating resources based on knowledge derived from FCM, organizations can anticipate several potential financial benefits because this tool identifies the need for new products or services or to expand into new markets, or to improve operational processes, among other things, which allow them to be in an excellent position with respect to their competitors.

6 Conclusions

The use of computational techniques to analyze factors and assess innovation helps diagnose and assess innovation management in an organization. This study proposes a computational tool for analyzing the key variables involved in innovation processes: types of innovation, innovative activities, resources, and external contexts. In addition, the model allows diagnosing innovation based on the factors and evaluation of innovation. The diagnostics allow knowing the current level of innovation in the company.

Thus, this work proposed an FCM for the evaluation of the level of innovation in an organization, which can be applied in different contexts, as evidenced by the Colombian case studies and the global case study, with very good results. This is the first work to propose an FCM for innovation evaluation using innovation types, innovative activities, resources and external contexts, such as the macro variables of the organizational environment that influence internal competence development, based on the Oslo Handbook. The model developed achieved an accuracy of 80% for Colombian cases and 92% for global cases for evaluating the innovation levels in the organizations.

The FCM allowed modeling this problem using highly flexible classification patterns. Therefore, the obtained FCM is very useful in evaluating the innovation of firms. The study found that the most influential factor is innovative activities. Innovative activities facilitate the development of innovations necessary for the proper functioning of the company. Their lack leads to the fact that organizations waste time, slow down processes, and do not pay attention to market changes. Thus, the model developed allows assessing the impact of innovation activities and other factors in innovation evaluation.

For future work, it is important to automate the definition of the relationships between the concepts of the context and of our FCM. Also, in the event that a concept does not exist in the FCM but there is data, it is important to add new concepts and verify if the behavior of the FCM improves or gets worse. It is also possible to improve the behavior of the FCM by analyzing the variables with a greater impact on the level of innovation and very dynamic in the context since they are the variables to be observed in real-time to follow the evolution of the innovative behavior of an organization. Finally, a future work must define concrete measures to improve innovation and enhance a culture of innovation within an organization based on the causal relationships determined by our FCM.

Competing Interests

None reported.

Funding Information

Ana Gissel Gutiérrez Buitrago is supported by a PhD grant financed by Universidad EAFIT. All the authors would like to thank the "Vicerrectoría de Descubrimiento y Creación", of Universidad EAFIT, for their support on this research.

Author contribution

A Gutiérrez: Methodology Formal analysis, Experiments, codification, Writing; *J. Aguilar:* Conceptualization, Methodology Formal analysis, Simulation, Writing, *A. Ortega:* Formal analysis, Writing; *E. Montoya:* Formal analysis, Writing, Funding acquisition.

Data Availability Statement

The data used will be available upon request.

References

[1] Oslo Manual, The Measurement of Scientific, Technological and Innovation Activities, 2018. OECD, doi: 10.1787/9789264304604-en. Available in: https://www.oecd-ilibrary.org/docserver/9789264304604-en.pdf?expires=1671972973&id=id&accname=guest&checksum=E032C1868203285E2094F14D22DD1EC9

[2] H. Jaramillo, "Normalización de Indicadores de Innovación Tecnológica en América Latina y el Caribe" RICYT / OEA / CYTED, 2021. Available in: https://minciencias.gov.co/sites/default/files/bogota.pdf

[3] B. Kosko, "Fuzzy cognitive maps," International Journal of Man-Machine Studies. Vol. 24, pp. 65–75, 1986.

[4] B. Kosko, Fuzzy Engineering. Phi Learning, 201997.

[5] J. Aguilar, "A fuzzy cognitive map based on the random neural model. Lecture Notes in Artificial Intelligence, vol. 2070. pp 333–338, 2001.

[6] J. Aguilar, J. Contreras, "The FCM Designer Tool". In: Glykas, M. (eds) Fuzzy Cognitive Maps. Studies in Fuzziness and Soft Computing, vol. 247. pp 71–87 Springer, 2010.

[7] Z. Wei, L. Lu, and Z. Yanchun, "Using fuzzy cognitive time maps for modeling and evaluating trust dynamics in the virtual enterprises," Expert Syst Appl, vol. 35, no. 4, pp. 1583–1592, 2008, doi: 10.1016/j.eswa.2007.08.071.

[8] K. Kok, "The potential of Fuzzy Cognitive Maps for semi-quantitative scenario development, with an example from Brazil," Global Environmental Change, vol. 19, no. 1, pp. 122–133, 2009, doi: 10.1016/j.gloenvcha.2008.08.003.

[9] Bhanu. Prasad, "Soft computing applications in business". Springer, 2008.

[10] H. Sánchez, J. Aguilar, O. Terán, J. Gutiérrez de Mesa, "Modeling the process of shaping the public opinion through Multilevel Fuzzy Cognitive Maps", Applied Soft Computing, Vol. 85, 2019, doi: 10.1016/j.asoc.2019.105756.

[11] J. Aguilar, "Different dynamic causal relationship approaches for cognitive maps," Applied Soft Computing Journal, vol. 13, no. 1, pp. 271–282, 2013, doi: 10.1016/j.asoc.2012.08.037.

[12] J. Aguilar, "Multilayer Cognitive Maps in the Resolution of Problems using the FCM Designer Tool," Applied Artificial Intelligence, vol. 30, no. 7, pp. 720–743, Aug. 2016, doi: 10.1080/08839514.2016.1214422.

[13] M. Ladeira, F. Ferreira, J. Ferreira, W. Fang, P. Falcão, aA. Rosa, "Exploring the determinants of digital entrepreneurship using fuzzy cognitive maps," International Entrepreneurship and Management Journal, vol. 15, no. 4, pp. 1077–1101, 2019, doi: 10.1007/s11365-019-00574-9.

[14] E. Gomez, H. Guerrero "Design of a fuzzy cognitive map as an approach to a contemporary socioeconomic structure: Contributions from institutionalism," Análisis Económico, vol.35, pp.65-89. 2020

[15] A. Infante-Moro, J. C. Infante-Moro, and J. Gallardo-Pérez, "Fuzzy cognitive maps and their application in social science research: A study of their main problems," Education in the Knowledge Society, vol. 22. 2021. doi: 10.14201/EKS.26380.

[16] W. Froelich, "Towards improving the efficiency of the fuzzy cognitive map classifier," Neurocomputing, vol. 232, pp. 83–93, 2017, doi: 10.1016/j.neucom.2016.11.059.

[17] A. Amirkhani, E. I. Papageorgiou, A. Mohseni, M. Mosavi, "A review of fuzzy cognitive maps in medicine: Taxonomy, methods, and applications," Computer Methods and Programs in Biomedicine, vol. 142. pp. 129–145, 2017. doi: 10.1016/j.cmpb.2017.02.021.

[18] E. Puerto, J. Aguilar, C. López, and D. Chávez, "Using Multilayer Fuzzy Cognitive Maps to diagnose Autism Spectrum Disorder," Applied Soft Computing Journal, vol. 75, pp. 58–71, 2019, doi: 10.1016/j.asoc.2018.10.034.

[19] F. Lucila, D. Pozos, M. Acosta Márquez, "Analysis and relevance in business development." Pensamiento & Gestión, Vol. 40, pp. 184-202, 2016

[20] D. Weldeslasie, M. Ahmed, "The application of expert systems in manufacturing sector," International Journal of Innovative Technology and Exploring Engineering, vol. 9, pp. 4744–4749, 2019, doi: 10.35940/ijitee.A4381.119119.

[21] L. Ferrer, C. Karina, P. González, I. Lilia, M. Mendoza, "Innovation as a key factor for improving the competitiveness of SMES in the department of Atlantic, Colombia," Dictamen Libre, vol. 16, pp. 21–36, 2015.

[22] M. Sánchez. V. Cervantes . P. Peralta, "Innovation Management in Small and Medium Enterprises of Barranquilla- Colombia", Revista de Ciencias Sociales, vol. 22, pp. 78-91, 2016

[23] R. Axelrod, "Structure of decision: The cognitive maps of political elites," Princeton University Press, 1976.

[24] J. Salmeron, S. Rahimi, A. Navali, A. Sadeghpour, "Medical diagnosis of Rheumatoid Arthritis using data driven PSO–FCM with scarce datasets," Neurocomputing, vol. 232, pp. 104–112, 2017, doi: 10.1016/j.neucom.2016.09.113.

[25] H Fontalvo, F. Calvano, M. Puente, E. Arias, B. Fontecha, "Innovation in the organizational development of small and medium-sized food and beverage companies in Colombia". International Journal of Entrepreneurship, vol. 25, 2021.

[26] M. Navarro-Caballero, L. Hernández-Fernández, E., Navarro-Manotas, J. Hernández-Chacín, "Innovation in micro, small and medium-sized family business in the manufacturing sector in Atlántico-Colombia. Revista De Ciencias Sociales, vol. 26, pp. 124-144, 2020. Doi: 10.31876/rcs.v26i4.34653

[27] F. Khemakhem, H. Ellouzi, H. Ltifi and M. B. Ayed, "Agent-Based Intelligent Decision Support Systems: A Systematic Review", IEEE Transactions on Cognitive and Developmental Systems, vol. 14, pp. 20-34, 2022, doi: 10.1109/TCDS.2020.3030571.

[28] Ruta N: innovation and business center of Medellín "Raw data on Innovation in Valle de Aburrá organizations", 2020, https://www.rutanmedellin.org/

[29] "Global Innovation Index 2022" dataset, World Intellectual Property Organization, 2022. Available in: https://www.wipo.int/publications/es/details.jsp?id=4622&plang=EN