

Digital Object Identifier: 10.24054/rcta.v2i46.3530

Multi-agent system focused on distributed artificial intelligence processes for order capture and routing

Sistema multiagente enfocado en procesos de inteligencia artificial distribuida para la captura y el ruteo de pedidos

MSc. Anny Astrid Espitia Cubillos ^[b], PhD. Robinson Jiménez-Moreno ^{[b2}, Ing. Mateo Andrés Pulido Aponte ^[b]

¹Universidad Militar Nueva Granada, Facultad de Ingeniería, Programa de Ingeniería Industrial, Bogotá, Colombia ²Universidad Militar Nueva Granada, Facultad de Ingeniería, Programa de Ingeniería Mecatrónica, Bogotá, Colombia.

Correspondence: anny.espitia@unimilitar.edu.co

Received: january 20, 2025. Accepted: june 21, 2025. Published: july 01, 2025.

How to cite: A. A. Espitia Cubillos, R. Jiménez Moreno, and M. A. Pulido Aponte, "Multi-agent system focused on distributed artificial intelligence processes for order capture and routing", RCTA, vol. 2, no. 46, pp. 13–21, Jul. 2025. Recovered from <u>https://ojs.unipamplona.edu.co/index.php/rcta/article/view/3530</u>

> This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.



Abstract: This paper presents the design of a multi-agent system focused on text-operated distributed artificial intelligence processes that works following a dialog model, oriented to the efficient management of industrial product orders. This system integrates five reasoning and action agents: the general agent powered by Google's Gemini artificial intelligence language model that carries the conversation flow and redirects the tasks to the other agents, an agent that identifies customers, an agent that recognizes the products, an agent that generates the orders and an agent that routes. The multi-agent system is designed to interact virtually and conversationally with users, facilitating order creation and management through an innovative approach that integrates advanced natural language processing technologies, vector and relational databases, and optimization methods. Finally, for its validation in a real environment, data from a company that produces and distributes cleaning products are used, allowing the development of different tests for the identification of customers, products, quantities and the establishment of order delivery routes. It is concluded that the conversational flow and techniques used allow users to make queries about customers with an average accuracy of 77.38% and about products with an average accuracy of 88.57%, even in scenarios with semantic ambiguities, managing orders in an intuitive way. It is also possible to optimize order routing by simultaneously considering two criteria that can be weighted: customer importance and distance traveled.

Keywords: app, reasoning and action agent, optimization, databases, routing, artificial intelligence, multi-agent system.

Resumen: Este documento presenta el diseño de un sistema multiagente enfocado en procesos de inteligencia artificial distribuida operado por texto que funciona siguiendo un modelo de dialogo, orientado a la gestión eficiente de órdenes de pedido de productos industriales, éste sistema integra cinco agentes de razonamiento y acción: el agente general potenciado con el modelo de lenguaje de inteligencia artificial Gemini de Google que lleva



el flujo de la conversación y redirecciona las tareas a los otros agentes, un agente que identifica clientes, un agente que reconoce los productos, un agente que genera los pedidos y un agente que rutea. El sistema multiagente está diseñado para interactuar de manera virtual y conversacional con los usuarios, facilitando la creación y gestión de pedidos mediante un enfoque innovador que integra tecnologías avanzadas de procesamiento natural del lenguaje, bases de datos vectoriales y relacionales, y métodos de optimización. Finalmente, para su validación en un entorno real, se utilizan datos de una empresa de producción y distribución de productos de aseo, lo que permite desarrollar distintas pruebas para la identificación de clientes, productos, cantidades y establecimiento de rutas de entrega de pedidos. Se concluye que el flujo conversacional y técnicas utilizadas permite a los usuarios realizar consultas sobre clientes con una exactitud promedio del 77,38% y de productos con una precisión promedio del 88.57%, aún en escenarios con ambigüedades semánticas, logrando gestionar órdenes de manera intuitiva. Se logra además optimizar el ruteo de los pedidos considerando simultáneamente dos criterios que pueden ponderarse: importancia del cliente y distancia recorrida.

Palabras clave: aplicación, agente de razonamiento y acción, optimización, bases de datos, ruteo, inteligencia artificial, sistema multiagente.

1. INTRODUCTION

In the industrial sector, companies need to operate efficiently, which requires the continuous improvement of processes including the use of technologies that reduce the workload of employees and allow them to concentrate on the tasks that add value and improve customer responsiveness, properly addressing these two challenges are a factor of business success.

Information technology systems were developed to respond to the increasing complexity of processes [1]. Chatbots as technological systems enable the communication of people and are constituted as an enterprise user interface. Chat agents are an evolution of chatbots.

Currently, chatbots are used in industry for data entry in production [2], which can be integrated with artificial intelligence to gather information quickly to facilitate and support decision making in manufacturing [3]. In logistics, they are one of the fundamental technologies applied for humanmachine interaction [4]. Also in the service sector, chatbots help users to schedule appointments, make reservations and set reminders [5], in general, they allow to guide and help internal (workers) and external (customers) users [6].

Industry 4.0 seeks to obtain intelligent products, through manufacturing, supply chain management and customer handling, aspects developed by intelligent systems. In addition, Industry 5.0. involves the use of human-centered systems [7], which has led organizations to seek the successful integration of digital technologies with people.

Therefore, it is possible to understand a multi-agent system focused on distributed artificial intelligence processes [8] as a digital tool that facilitates the relationship with the customer as in the application case of this article, where it is used for virtual capture and order routing. Or as proposed in [9] as a tool that helps workers to perform tasks and make decisions. A traditional chatbot responds to a tree structure, which guides an interaction by numerical responses or simple yes or no answers.

In [10], a chat-like logistics agent is designed to manage material transportation that is responsible for delivering information to the programmer responsible for choosing the right vehicle. Also in [11] a design of a manufacturing operations management system is presented that includes an artificial intelligence supported chatbot interface with a prediction system that retrieves live information from the production database.

However, [5] postulates that by using chatbots and virtual assistants, science and technology innovation developments have the potential to supplant employees. But the approach presented here coincides with that of [12] in the sense of rescuing the potential of collaborative work between humans and AI, within the framework of Industry 5.0 thinking in a human-centered development.

Large language models (LLMs) work well on general purpose tasks, but fail in specific areas that are updated. In [13] they integrate retrieval-

augmented generation (RAG) with reasoning and action agents (ReAct) in order to avoid failures in LLMs when generating biomimetic designs.

Based on the analysis of the state of the art, this article aims to present the design of a multi-agent system focused on artificial intelligence processes whose contribution consists of automatically guiding the taking of products and their routing for the dispatch of orders, based only on customer-agent interaction. Its validation is carried out through the application to a company that produces and distributes cleaning products.

The document is structured in five sections: the first one corresponds to the introduction, the second part presents the methodology, the third one describes the development, the fourth one reports the main results and the last one presents the conclusions reached.

2. METHODOLOGY

For the development of the multi-agent system focused on distributed artificial intelligence processes that allows the capture and subsequent routing of orders, five phases were defined.

In the first phase, the reasoning and action agent (Re-Act) [14] is selected which, in order to reduce possible failures in the LLM, requests the user to choose an action from a list of options. The agent is powered by Google's Gemini artificial intelligence model [15], which provides advanced capabilities to manage complex conversational flows, improving the user experience in creating order requests.

In the second phase, vector databases [16] are used for two purposes: to perform efficient searches of customer names and to identify products. In this case using the open source vector database Chroma [17], vector representations of customer names and product names and descriptions are stored, allowing approximate searches that are robust to misspellings or variations in the query. Additionally, in the case of products, it allows suggestions of similar items to make it easier for the user to select the option of his preference. The vector database approach improves the accuracy and efficiency of searches, overcoming the limitations of traditional systems based on exact text.

In the third phase, PostgreSQL [18] is used to manage relational databases of each product's inventory, orders, customers and even the logistics routes associated with the system. This component ensures the integrity and consistency of operational information. PostgreSQL enables the operation of the relational system for:

- Inventory management: Record and monitor the status of products.
- Customer and order control: Maintain a structured and reliable history.
- Route calculation: Integrate distance matrix and purchase orders to generate optimized routes.

LlamaIndex [19] is used to integrate vector and relational databases with the Gemini AI model [15], which translates user queries into concrete actions, such as searching for products, checking inventories or querying customer data, in a fast and efficient way.

In the fourth phase, route optimization is carried out since the delivery of orders depends on the definition of these routes, for which a distance matrix summarizing the possible connections between origin and destination points is used. The route calculation simultaneously considers two criteria that can be weighted: customer importance (order priority) and distance traveled, using Kohonen maps [20] to find minimum cost solutions. This approach ensures the selection of optimal routes, which maximize customer satisfaction according to their relevance, minimize delivery times, or simultaneously satisfy these two objectives in a weighted manner.

Finally, in the fifth phase, the Streamlit framework [21] is used to create a user-friendly interface that allows users to easily interact with the conversational multi-agent system, visualize relevant data, such as product recommendations, logistic routes, and estimated delivery times. In addition to displaying in real time the status of inventories, purchase orders and optimal routes generated.

3. DEVELOPMENT

According to the European Commission's AI Law low-risk chatbots are required to comply with transparency obligations [22], therefore the development of the solution focused on the integration of multiple advanced technologies to build a robust and efficient conversational system, covered in the framework of the mentioned law. As mentioned, a ReAct agent [14] powered by Google's Gemini language model [15] was implemented, orienting its implementation to natural interaction with users to manage order requests. The designed multi-agent system integrates five Re-Act agents [14]: the general agent powered with Google's Gemini artificial intelligence language model that carries the conversation flow and redirects tasks to the other agents, an agent that identifies customers using the Chroma vector database [17], an agent that recognizes products also using the Chroma vector database, an agent that manages relational databases of the inventory using PostgreSQL [18] to generate orders and keep history, and an agent that routes using Kohonen maps [20]. Streamlit is used to create the interface with the users. The structuring of the multi-agent system and the various tools that support its operation are presented in Fig. 1.

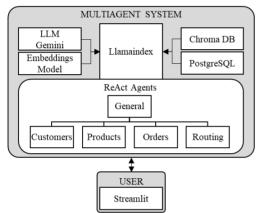


Fig. 1. Structure of the multi-agent system design

The multi-agent system focused on distributed artificial intelligence processes for order capture and routing works based on the flowchart shown in Fig. 2. The process represented in the flowchart allows the multi-agent to manage the direction of the conversation completely, evaluating the data and information available to redirect it if necessary and request the relevant information to the user.

Regarding the vector databases using Chroma [17], a preprocessing was performed mainly on the customer data, eliminating recurring words such as: Conjunto, Edificio, Multifamiliar, etc, with the objective of improving the results when performing semantic searches with user input. The same preprocessing was performed with the product data.

For its implementation it was required to store the data of products, customers, distances, orders, and routes in relational databases of the validation environment, to maintain accuracy and proper handling of the data. The fields that contain the tables in each of the databases and their type are shown in Fig. 3.



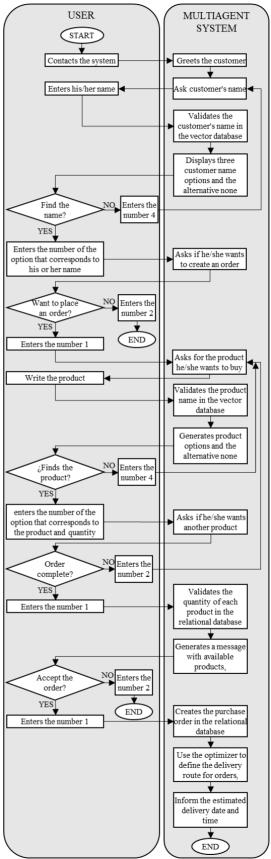


Fig. 2. Multi-agent system flowchart

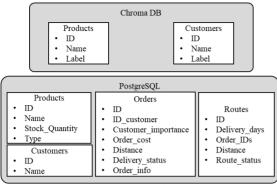


Fig. 3. Database information

The combination of relational and vector data allows the multi-agent system focused on distributed artificial intelligence processes for order capture and routing to be versatile and effective in decision making.

The calculation of routes was performed by using the Kohonen map algorithm [20] due to its competitive and cooperative learning approach forming relationships between the distances of the sites to define the route, also through the normalization of distances another weight factor can be included for the efficient calculation of routes, such as importance, leaving a variable weighting value from 0 to 1 between the distance criterion and the importance criterion to evaluate the results obtained. The 10 equations used for distance calculation are presented below.

• Dynamic input vector:

$$S_0 = [0, d_{01}, d_{02}, \dots, d_{0n}]$$

$$S_1 = [d_{10}, 0, d_{12}, \dots, d_{1n}]$$

$$S_2 = [d_{20}, d_{21}, 0, \dots, d_{2n}]$$

$$S_n = [d_{n0}, d_{n1}, d_{n2}, \dots, 0]$$

Where S_i represents the vector that groups the data of the distances (d_{ij}) from origin i to destination j.

• Dynamic weight matrix:

$$M = \begin{bmatrix} w_{00} & w_{01} & \cdots & w_{0n} \\ \vdots & \ddots & \vdots \\ w_{n0} & w_{n1} & \cdots & w_{nn} \end{bmatrix}$$

Where W_{ij} corresponds to the weight from origin i to destination j

• Distance scaling including the importance factor.

$$d_{s} = \left(\frac{d_{s} - d_{min}}{d_{max} - d_{min}} \times W_{d}\right) + \left(\frac{I_{s} - I_{min}}{I_{max} - I_{min}} \times W_{i}\right)$$

Where W_d represents the weighting of the distance traveled criterion and W_i represents the weighting of the importance of the client, it is important to bear in mind that it must be guaranteed that in all cases the following criteria are met:

$$\sum W_d + W_i = 1$$

• Euclidean distance:

$$d_{euc} = \sqrt[2]{\sum (x_{i0} - x_{i1})^2}$$

Where X_{ij} represents the location of each customer.

Best Coincident Unit BCU

$$BCU = min([d_{euc0}, d_{euc1}, \dots, d_{eucn}])$$

• Weight update radius:

$$r(t) = r_o e^{-\frac{t}{\lambda}}$$

Where t representas the time periodo.

• Product factor:

$$\beta_t = e^{-\frac{d^2}{2r_t^2}}$$

• Learning rates:

$$\lambda_0 = \frac{T}{\ln (ro)} \approx 0.5$$
$$\lambda(t) = \lambda_0 e^{-\frac{t}{1000}}$$

• Updating of weights:

$$M_t = M_{t-1} + \lambda_t \times \beta_t (d_{si} - M_{t-1})$$

University of Pamplona I. I. D. T. A Fig. 4. shows the design of the interface developed from the point of view of the client (external user).



Fig. 4. User interface

Finally, Fig. 5 shows the record of operations that can be consulted in the interface by the commercial advisor (internal user).

		Resumen					
vedades							
Invento	nrio						
COD_PRODU	CTO NOMBRE		TIPO_PR	ористо	PRECIO_PRI	DDUCTO	STOK
1	0105 Limpiador Dirt Free Office Crema po	410 g	в			29850	10
1	0117 Limpiador Dirt Free Office Crema po	Galon	в			99696	10
1	0118 Limpiador Dirt Free Office Crema po	1/2 Galon	в			91300	10
1	0202 Limpiador Dirt Free Office Liquido p	or 120 cc	в			161576	10
1	0212 Limpiador Dirt Free Office Liquido p	or 5 Galones	٨			299015	10
	S NOMBRE	IMPO	RTANCIA	DISTANC		7100 01	IENTE
190	TSG The It Experts Principal						
190		4.	6		IA_ORIGEN 14 12	вс	
	Agrupacion De Vivienda La Esperanza P.	۹.	6		14	BC	
164	Agrupacion De Vivienda La Esperanza P. Edificio Balmoral II P.H.		6 8		14	BC CB	
164	Agrupacion De Vivienda La Esperanza P.I Edificio Balmoral II P.H. Edificio Kasaury Aptos-Primera Etapa-P.I	4.	6 8 7		14 12 8	BC CB CA	
164 199	Agrupacion De Vivienda La Esperanza P.I Edificio Balmoral II P.H. Edificio Kasaury Aptos-Primera Etapa-P.I Coniunto Residencial Aticos del Salitre P	4.	6 8 7 6		14 12 8 16	BC CB CA BC	
164 199 164 166 Pedido	Agrupacion De Vivienda La Esperanza P.I Edificio Balmoral II P.H. Edificio Kasaury Aptos-Primera Etapa-P.I Coniunto Residencial Aticos del Salitre P	4. H.	6 8 7 6 9		14 12 8 16 25	BC CB CA BC CC	
164 199 164 166 Pedido id_pedido	Agrupacion De Vivienda La Esperanza Pi Edificio Balmonal II P.H. Edificio Kasaury Aptos-Primera Etapa-Pi Coniunto Residencial Aticos del Salitre P S	4. H.	6 8 7 6 9		14 12 8 16 25	BC CB CA BC CC	

4. RESULTS

4.1. Customer identification validation

For the validation of customer identification, a database with customer information was generated and modified to include three scenarios: one with complete and correct information, one with incomplete information and one with incorrectly typed information, corresponding to the inaccuracies that the designers consider may be present on the part of the customers.

After performing the semantic search of the name of different customers in the multi-agent system focused on distributed artificial intelligence processes for order capture and routing, the results presented in Table 1 are obtained.

Table 1: 1	Results d	of customer	identification tests

Information entered	Percentage of success
Full name	100,00%
Incomplete name	57,14%
Incorrectly typed name	75,00%
Average	77,38%

Table 1 shows that the identification of clients in the first scenario obtained the expected result, which corresponds to a 100% success rate. When performing the semantic search for customer names that were not complete, 57.14% of the cases were correct. While the semantic search for customer names entered by the user with typing errors yielded a 75% accuracy rate. This allows calculating a hit percentage by simple average of the three scenarios evaluated in the test of 77.38%.

To improve the results, it is advisable to better label the customer names, use a word embedding model [23] with higher dimensions and/or reinforcement learning [24] for training the artificial intelligence.

4.2. Product identification validation

For the validation of product identification, the same procedure was applied as for customer identification, in which the database with product information was modified to have the same three scenarios: one with complete and correct information, one with incomplete information and one with incorrectly typed information. Table 2 shows the results obtained for the product tests in the three scenarios considered.

Table 2: Product identification test results

Information entered	Percentage of success		
Full name	100,00%		
Incomplete name	65,71%		
Incorrectly typed name	100,00%		
Average	88,57%		

As shown in Table 2, in the first scenario a 100% success rate was again achieved, and the same result was also achieved when the user entered the data incorrectly. There were only failures when the product information was not complete; in this scenario a 65.71% success rate was obtained. The percentage of success by simple average of the three scenarios evaluated in the test was 88.57%.



It can also be inferred that better labeling and a model with larger word embedding dimensions [23] could improve the accuracy of identification when the information entered by the user is incomplete, in addition it is possible to use RAG techniques [13] to perform several searches by synonym of the products.

4.2. Route calculation validation

To validate the calculation of the optimal routes, six scenarios were considered using Kohonen maps with different weightings of the two decision criteria established: importance of the client and distance traveled.)

Table 3 shows the results obtained from the route calculation.

Table 3: Average distance traveled using different weightings	
for the criteria	

Weigh	Distance traveled (Km)				
Importance	Distance	Day 1	Day 2	Day 3	Average
0%	100%	78,90	88,94	152,55	106,795
100%	0%	96,36	101,27	199,24	132,288
60%	40%	63,99	84,47	153,22	100,560
40%	60%	92,13	91,23	153,98	112,444
80%	20%	96,66	82,29	153,22	110,724
20%	80%	78,90	78,68	156,44	104,672

As can be seen in Table 3, the scenario with the best results is obtained when a weighting of 40% is given to the total distance traveled and 60% to the importance of the customer, which presents lower distance values and positions the most important sites at the beginning of the route.

Fig. 6. presents an example of the initial route without optimization, the numbers indicate the order of delivery of the orders, it covers a total distance of 26.4 km.



Fig. 6. Initial route

In Fig. 7, images are included showing how, during some of the 5000 epochs of training, the neurons self-organize at the different delivery sites taking

into account the distances with longitude and latitude coordinates.

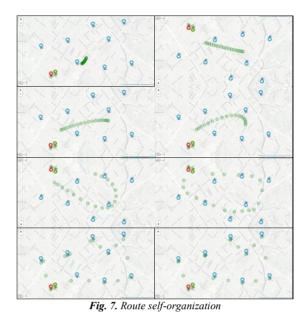


Fig. 8 shows the optimized route with a total distance of 17.1 km, with a 35.23% reduction in travel time.



Fig. 8. Optimized final route

It is important to note that in Fig. 6 to 8, there is an error of up to 1 Km in the location, this margin of error is given because the distances are calculated using the latitudes and longitudes of each delivery site, but the real distance that would be obtained by knowing exactly the streets that allow reaching the delivery points is not taken into account.

5. CONCLUSIONS

It was found that vector databases allow identifying customers in an acceptable manner (77.38%) even in scenarios with ambiguities, i.e. the development operates with a margin of errors in the search typing, which is favorable. They also allow identifying products in an adequate manner (88.57%), which serves to recommend similar options to the user according to preferences. Chroma improves the user

University of Pamplona I. I. D. T. A experience by seeking accuracy and speed in searches, even in imprecise or poorly worded queries.

The integration of advanced natural language processing technologies, vector and relational databases, and route optimization methods provided a multi-agent system design focused on distributed artificial intelligence processes that is considered robust and efficient for order management in the tests performed. In turn, it was shown that it is able to adapt to user needs and optimize logistics processes, improving the customer experience and reducing operating costs associated with the taking and final delivery of orders.

As future work, it is proposed to consider voice interaction to facilitate the conversation between the user and the intelligent chat agent. Improve the identification of customers and products when the information entered by the customer is incomplete, by techniques such as reinforcement learning, word embeddings and RAG. Also keep in mind other optimization criteria for establishing order delivery routes.

ACKNOWLEDGMENTS

Product derived from the research project entitled "Mejoramiento de procesos de control de inventario y alistamiento de pedidos en el marco de la industria 4.0" INV-ING-3970 funded by the Vice Rector's Office of Research of the Universidad Militar Nueva Granada, 2024.

REFERENCES

- A. Börold, D. Schweers y M. Freitag, «Towards Multimodal Information Systems for Assisting Humans in Production and Logistics Processes,» *Procedia CIRP*, vol. 120, pp. 1089-1094, 2023. https://doi.org/10.1016/j.procir.2023.09.130.
- [2] A. Dörner, P. Gerald, S. Michal y B. Marek, «Use of Lean Management Methods based on Eye-Tracking Information to make User Interfaces in Production more Humancentered,» vol. 128, pp. 514-519, 2024. https://doi.org/10.1016/j.procir.2024.04.014.
- [3] S. Vernim, H. Bauer, E. Rauch, M. T. Ziegler y S. Umbrello, «A value sensitive design approach for designing AI-based worker assistance systems in manufacturing,,» *Procedia Computer Science*, vol. 200, pp.

505-516, 2022. https://doi.org/10.1016/j.procs.2022.01.248.

- [4] M. Pellicelli, «Chapter Two A long road to maximizing efficiency,» de *The Digital Transformation of Supply Chain Management*, Elsevier, 2023. https://doi.org/10.1016/B978-0-323-85532-7.00009-8, pp. 37-61.
- [5] I. Ghannouchi, «Examining the dynamic nexus between industry 4.0 technologies and sustainable economy: New insights from empirical evidence using GMM estimator across 20 OECD nations,» *Technology in Society*, vol. 75, p. 102408, 2023. https://doi.org/10.1016/j.techsoc.2023.10240 8
- [6] A. Salis, A. Marguglio, G. D. Luca, S. Razzetti, W. Quadrini y S. Gusmeroli, «An Edge-Cloud based Reference Architecture to support cognitive solutions in Process Industry,» *Procedia Computer Science*, vol. 217, pp. 20-30, 2023. https://doi.org/10.1016/j.procs.2022.12.198.
- T. A. Aniceski, L. T. Pontes Miranda, O. Canciglieri Junior y G. Brittes Benitez, «The Four Smarts of Industry 4.0 and barriers for technology deployment: A TOE perspective,» *Computers & Industrial Engineering*, vol. 193, p. 110345, 2024. https://doi.org/10.1016/j.cie.2024.110345.
- [8] J. S. Dhatterwal, K. S. Kaswan y R. P. Ojha, «Chapter 13. The role of multiagent system in industry 4.0.,» de A Roadmap for Enabling Industry 4.0 by Artificial Intelligence, 2022. https://doi.org/10.1002/9781119905141.ch13 , pp. 227-246.
- [9] B. G. Mark, E. Rauch y D. T. Matt, «Systematic selection methodology for worker assistance systems in manufacturing,» *Computers & Industrial Engineering*, vol. 166, p. 107982, 2022. https://doi.org/10.1016/j.cie.2022.107982.
- [10] G. Bitsch, P. Senjic y J. Askin, «Dynamic adaption in cyber-physical production systems based on ontologies,» *Procedia Computer Science*, vol. 200, pp. 577-584, 2022.

https://doi.org/10.1016/j.procs.2022.01.255.

[11] S. Mantravadi, C. Møller, C. LI y R. Schnyder, «Design choices for nextgeneration IIoT-connected MES/MOM: An empirical study on smart factories,» *Robotics* and Computer-Integrated Manufacturing, vol. 73, p. 102225, 2022. https://doi.org/10.1016/j.rcim.2021.102225.

- [12] X. Sun y Y. Song, «Unlocking the Synergy: Increasing productivity through Human-AI collaboration in the industry 5.0 Era,» *Computers & Industrial Engineering*, vol. 200, p. 110657, 2025. https://doi.org/10.1016/j.cie.2024.110657.
- [13] C. Toukmaji y A. Tee, «Retrieval-Augmented Generation and LLM Agents for Biomimicry Design Solutions,» de *Proceedings of the AAAI Symposium Series*, 2024. https://doi.org/10.1609/aaaiss.v3i1.31210.
- [14] S. Yao, J. Zhao, D. Yu, N. Du, I. Shafran, K. Narasimhan y Y. Cao, «ReAct: Synergizing Reasoning and Acting in Language Models,» 2023.

https://doi.org/10.48550/arXiv.2210.03629.

- [15] H. R. Saeidnia, «Welcome to the Gemini era: Google DeepMind and the information industry,» *Library Hi Tech News*, 2023. https://doi.org/10.1108/LHTN-12-2023-0214.
- [16] Z. Jing, Y. Su y Y. Han, «When large language models meet vector databases: A survey,» 2024. https://doi.org/10.48550/arXiv.2402.01763.
- [17] P. N. Singh, S. Talasila y S. V. Banakar, «Analyzing Embedding Models for Embedding Vectors in Vector Databases,» de 2023 IEEE International Conference on ICT in Business Industry & Government (ICTBIG), Indore, India, 2023. https://doi.org/10.1109/ICTBIG59752.2023. 10455990.
- [18] A. Makris, K. Tserpes, G. Spiliopoulos, D. Zissis y D. Anagnostopoulos, «MongoDB Vs PostgreSQL: A comparative study on performance aspects,» *Geoinformatica*, vol. 25, p. 243–268, 2021. https://doi.org/10.1007/s10707-020-00407w.
- [19] R. K. Malviya, V. Javalkar y R. Malviya, «Scalability and Performance Benchmarking of LangChain, LlamaIndex, and Haystack for Enterprise AI Customer Support Systems,» de *IJGIS Fall of 2024 Conference. The New World Foundation.*, 2024. https://doi.org/10.21428/e90189c8.43aeb06e.
- [20] Q. Guan, X. Hong, W. Ke, L. Zhang, G. Sun y Y. Gong, «Kohonen Self-Organizing Map based Route Planning: A Revisit,» de 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS),

Prague, Czech Republic, 2021. https://doi.org/10.1109/IROS51168.2021.96 36025.

- [21] Y. Akkem, B. S. Kumar y A. Varanasi, «Streamlit Application for Advanced Ensemble Learning Methods in Crop Recommendation Systems – A Review and Implementation,» *Indian Journal of Science* and Technology, vol. 16, nº 48, pp. 4688-4702, 2023. https://doi.org/10.17485/IJST/v16i48.2850.
- [22] G. Schuh, M.-F. Stroh y L. Johanning, «Framework To Design Compliance Rules For Digital Technologies In Manufacturing Companies,» *Procedia CIRP*, vol. 119, pp. 103-108, 2023. https://doi.org/10.1016/j.procir.2023.04.005.
- [23] M. Habib, M. Faris, A. Alomari y H. Faris, «AltibbiVec: A Word Embedding Model for Medical and Health Applications in the Arabic Language,» *IEEE Access*, vol. 9, pp. 133875-133888, 2021. https://doi.org/10.1109/ACCESS.2021.3115 617.
- [24] Y. Matsuo, Y. LeCun, M. Sahani, D. Precup, D. Silver, M. Sugiyama, E. Uchibe y J. Morimoto, «Deep learning, reinforcement learning, and world models,» *Neural Networks*, vol. 152, pp. 267-275, 2022. https://doi.org/10.1016/j.neunet.2022.03.037.