Product Review Simulation using a Multi-Agent System

Fabián Suárez Maroto[1], Jaime Collado-Montañez[1,*] and Arturo Montejo-Ráez[1]

[1]Department of Computer Science (University of Jaén), Campus Las Lagunillas, s/n, Jaén, 23071, Spain

Abstract

This article describes a multi-agent system for simulating and analyzing product reviews. From a URL, it automatically extracts structured data, which is then evaluated by a user-configurable population of reviewer agents. These agents generate simulated reviews, which an analyst agent subsequently processes (sentiment, key aspects) and visualizes on a dashboard. The goal is to predict product reception and identify strengths/weaknesses before launch.

Keywords

Review simulation, multi-agent systems, web scraping, product evaluation

1. Introduction

The launch of new products into the market carries significant risks, with consumer acceptance being a critical factor for success. Understanding how a product will be perceived by different segments of the target audience before investing considerable resources in its production and marketing is of vital strategic importance. Traditionally, this understanding is obtained through market studies, surveys, or focus groups—methods that can be costly in terms of time and resources, and do not always dynamically capture the diversity of opinions.

In this context, we present a prototype of an agent-based modeling (ABM) simulation platform designed to automatically generate and analyze product reviews. Our system leverages the flexibility of multi-agent systems to model a heterogeneous population of virtual consumers, each with their own characteristics and evaluation criteria [1]. By providing structured product information, automatically extracted from the web, to this agent population, we can simulate the process of opinion formation and review generation.

The main objective is to offer product developers, marketing teams, and business strategists a tool to preliminarily assess the perceived strengths and weaknesses of a product for a specific target audience. This allows for identifying areas for improvement, adjusting features, or refining marketing messages before the official launch, thereby mitigating risks and optimizing market strategy.

This article describes the architecture and workflow of our prototype, detailing the roles of the different agents involved: the data extractor, the profile generation agent, the configurable reviewer agents, and the analyst agent. The underlying technologies and the platform's potential as a predictive analysis tool in the early stages of product development are discussed. Multi-agent systems based on large language models have proven particularly effective for complex tasks requiring coordination among multiple components [2].

This article is organized into the following main sections. First, Section 2 ("Related Work") reviews recent advances in opinion mining and highlights the emerging use of artificial personas for synthetic data generation and social simulations. Section 3 ("Architecture and Methodology") describes the structure of the multi-agent system, the workflow from data extraction to analysis, and the functions of each type of agent. Next, Section 4 ("Technological Framework") details the specific software tools and frameworks used for the prototype's implementation. Subsequently, Section 5 ("Application and

SEPLN 2025: 41st International Conference of the Spanish Society for Natural Language Processing, Zaragoza, Spain, 23-26 September 2025.

*Corresponding author.
fsm0037@red.ujaen.es (F. S. Maroto); jcollado@ujaen.es (J. Collado-Montañez); amontejo@ujaen.es (A. Montejo-Ráez)
0000-0002-9672-6740 (J. Collado-Montañez); 0000-0002-8643-2714 (A. Montejo-Ráez)

Benefits") explores the practical uses of the platform as a preliminary market research tool and the advantages it offers. Finally, Section 6 ("Conclusions and Future Work") summarizes the contributions of the work and outlines lines for future improvement and development of the system.

2. Related Work

Deep networks are also defining the state-of-the-art in opinion mining, like in aspect-based sentiment analysis [3] or fine-grained sentiment analysis [4], revitalizing a research topic that has been active for more than two decades.

In recent years, the capabilities of Large Language Models (LLMs) to emulate different human profiles has open a new topic focused on the use of *personas* as artificial profiles for many different tasks, mainly in the creation of synthetic data for model tranining [5] or agent personalization [6]. The use of personas has emerged as a new way to explore *in-silico* social research experimentation [7], as these simulations have been demostrated to be closely aligned with real-world population behaviour [8].

The application of artificial populations has not been yet explored in opinion mining, to the best of our knowledge, as a means to evaluate commercial products, offering a prospective information of potential weaknesses and strengths through artificial opinions. This approach could benefit from the maturity of opinion mining techniques to evaluate products from personas' opinions.

3. Architecture and Methodology

The proposed system follows a modular workflow implemented through a multi-agent architecture. Each type of agent has a specific responsibility within the overall simulation and analysis process, as illustrated in Figure 1.

The process begins with user intervention and unfolds through the following main stages: (1) Data input and information extraction, where product information is obtained; (2) Specification of parameters and generation of reviewer agents, where the simulated population is configured; (3) Generation of simulated reviews, where agents evaluate the product; and (4) Aggregated analysis and visualization, where results are processed and presented.

3.1. Data Input and Information Extraction

The user provides the URL of a product webpage of interest (Figure 2). The page is not required to belong to a specific domain; the system is designed to be generic. A first agent, the **Extractor Agent** (represented with a pickaxe in Figure 1), is responsible for performing web scraping on the provided URL.

This agent navigates the page, identifies, and extracts relevant product information such as name, description, technical features, price, and possibly images or key specifications [9]. The extracted information is structured and normalized into a standard format, like JSON, to facilitate its subsequent processing by other agents and the UI implementation (Figure 3). Optionally, the user could also provide this information directly in JSON format if they already have it or prefer to manually define the product features to be evaluated. Modern web scraping techniques allow for precise and efficient extraction of structured product data for subsequent analysis [10].

3.2. Configuration and Generation of Reviewer Agents

A crucial step in the simulation is the detailed configuration of the **Population of Reviewer Agents**. To facilitate this process, the system provides dedicated graphical interfaces that allow the user to define multiple facets of these virtual agents and their behavior when generating reviews. Initially, through the interface shown in Figure 4, the user establishes the general parameters of the simulated population and the desired characteristics for the reviews. The system enables the configuration of various aspects, including:

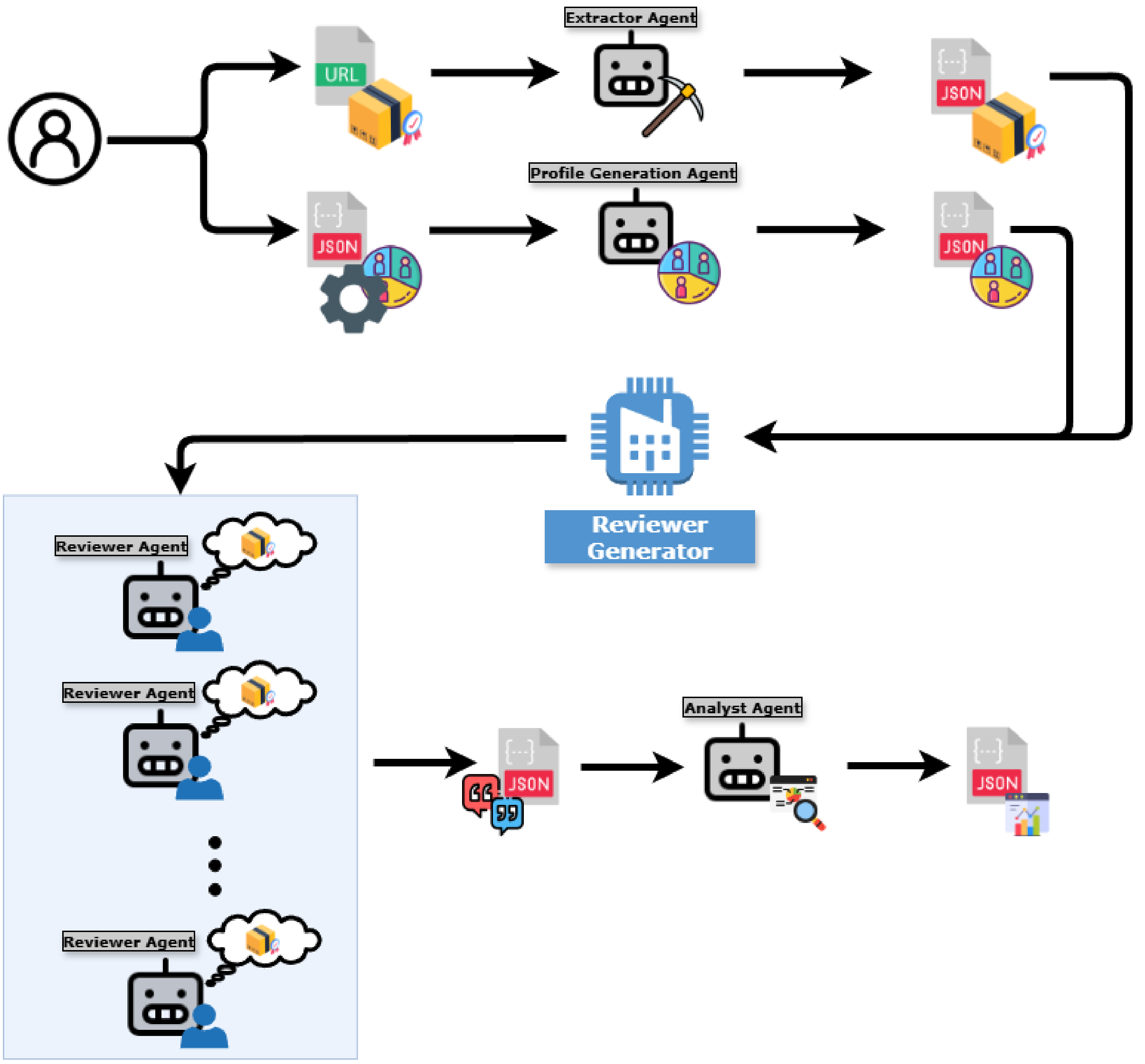

Figure 1: Architecture of the proposed multi-agent system. The roles of the Extractor Agent, the Population of Reviewer Agents, the Analyst Agent, and the information flow between them and the user are shown.

- **Population parameters:**
  - The total *size* of the population.
  - Demographic attributes such as *age* range, *education level* (allowing for specific or mixed profiles), and *gender* distribution.
- **Textual properties of the resulting reviews:**
  - *Positivity bias.*
  - *Verbosity* (length and elaboration).
  - The product-specific *level of detail* they should include.

For deeper modeling, the user can refine the **Personality Profile** of the agents using the interface in Figure 5. This allows adjusting values across various trait spectrums, defining tendencies on axes such as:

- *Introverted/Extroverted*
- *Independent/Cooperative*
- *Analytical/Creative*

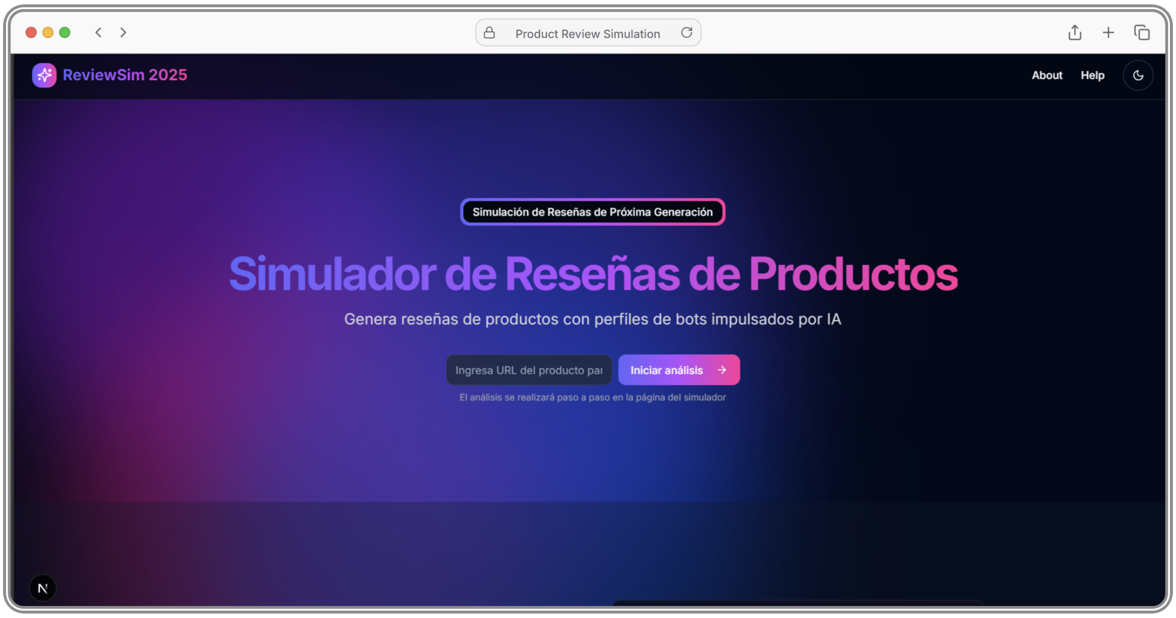

Figure 2: View of the main user interface page, allowing the user to enter a product URL to initiate its analysis or processing.

- *Ecologist/Non-ecologist*
- *Busy/Free time*
- *Safe/Risk-taker*
- *Disorganized/Organized*

The combination of these demographic, review style, and personality parameters gives the user the ability to model with great granularity the target audience whose reaction is to be simulated. These traits are expected to influence each agent's perspective, their points of focus or criticism, and the overall tone of the reviews they generate. Finally, once the user completes and confirms the configuration (e.g., using the "Generate bot profiles" button visible in Figure 5), the **Profile Generation Agent** uses the complete set of specifications to generate the reviewers' profiles Figure 6.

Upon user review of the generated profiles in the profile panel (Figure 6), the information is sent to the **Reviewer Generator** component (see Figure 1). Each profile is then used to instantiate a **Reviewer Agent**, which is ready to evaluate the product from its unique perspective.

3.3. Generation of Simulated Reviews

Once the structured product information (JSON) is available and the population of reviewer agents has been created, the product information is distributed to each agent in the population. Each individual **Reviewer Agent** (represented by the robot thinking about the product in Figure 1) processes the product information through the prism of its own defined parameters and personality. Using LLM models, each agent generates a simulated textual review (the prompt used is included in Appendix A). This review will reflect its particular evaluation of the product, highlighting positive or negative aspects based on its configured criteria [11]. The generated reviews are also collected in JSON format along with metadata from the agent that generated it, enabling their display on the interface (Figure 7). Recent advances in neural text generation have enabled the creation of more realistic and contextually relevant content, significantly improving the quality of simulated reviews [12].

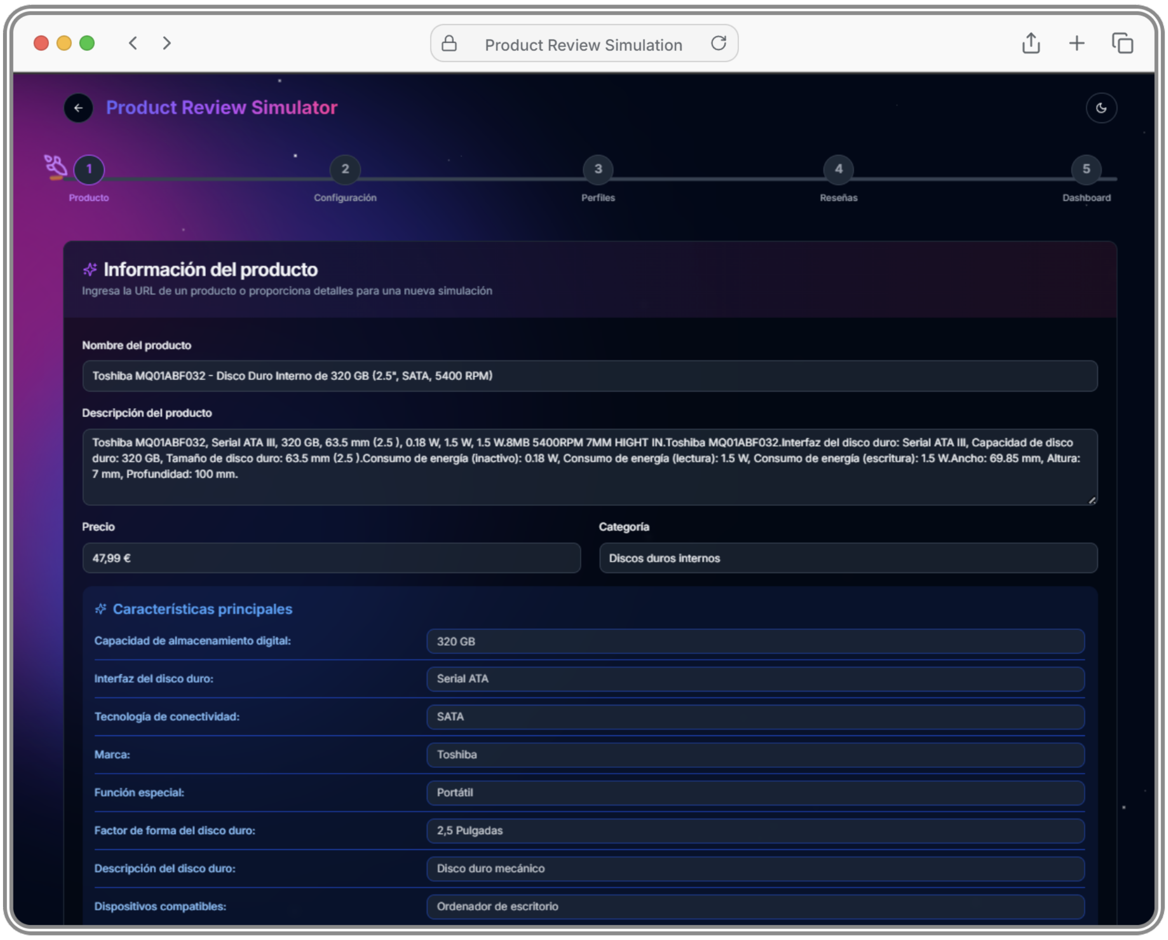

Figure 3: Visualization of the data collected by the Extractor Agent. The information presented is dynamic and varies according to the characteristics of the entered product.

3.4. Aggregated Analysis and Visualization

The set of simulated reviews generated by the population is sent to an **Analyst Agent** (represented by the robot with a magnifying glass over a web page in Figure 1). This agent is tasked with processing and analyzing all the reviews. Analysis techniques performed automatically by the LLM can include:

- Sentiment analysis: Determine the overall polarity (positive, negative, neutral) of the reviews [13].
- Aspect-Based Sentiment Analysis: Identify specific product features mentioned in the reviews and the sentiment associated with each [14].
- Topic identification: Group reviews or comments by recurring themes.
- Quantitative metrics: Calculate simulated average scores, rating distribution, frequency of feature mentions, etc.

The results of this analysis are consolidated and presented to the end-user through a **Dashboard** (represented by the JSON with graphs at the end). This panel (Figure 8) visualizes key findings, allowing the user to quickly and effectively understand the simulated product reception, identify patterns, and detect perceived strengths and weaknesses (Figure 9) by the simulated target audience. Evolved sentiment analysis techniques allow for a more nuanced understanding of the opinions expressed in reviews [15].

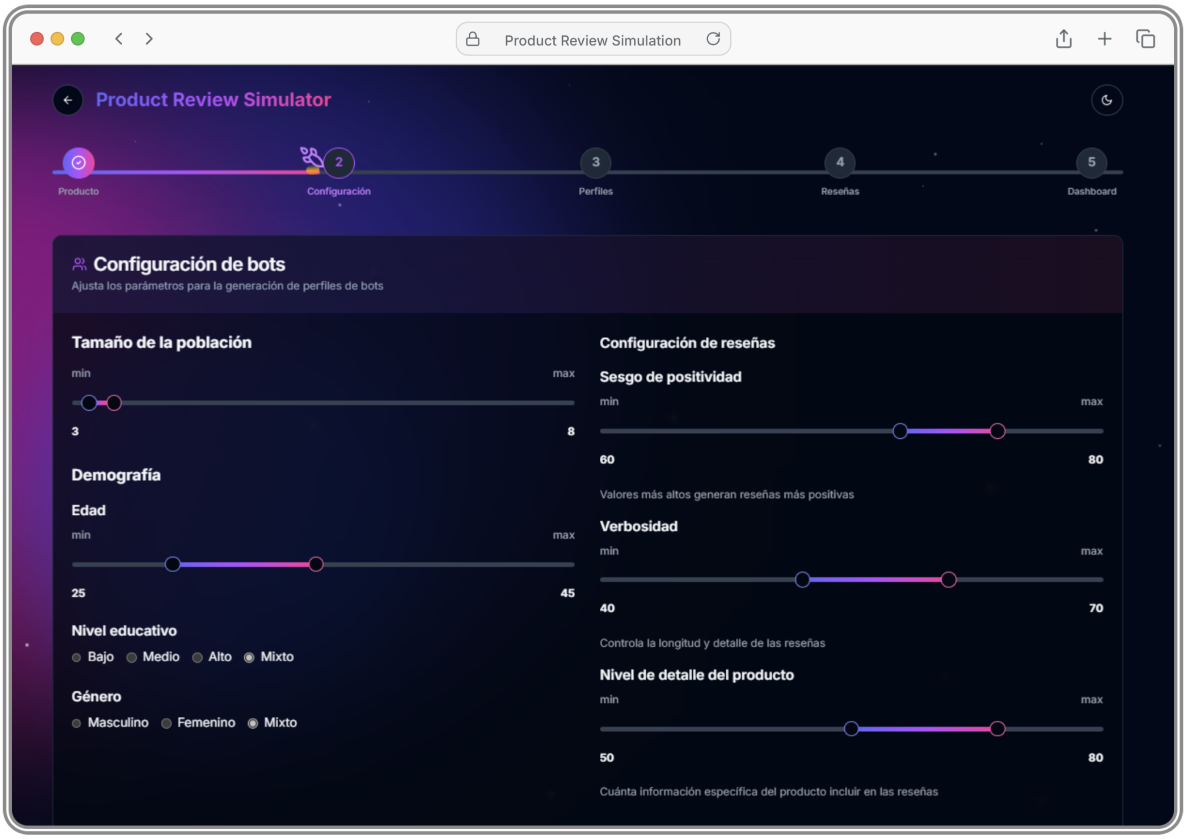

Figure 4: Interface for configuring demographic parameters of the agent population and general characteristics of the reviews to be generated.

4. Technological Framework

The implementation of the *Product Review Simulator* prototype is based on a set of technologies selected for their suitability for developing agent-based systems and interactive applications. The key components of the technology stack are as follows:

1. **Agent Orchestration Framework: CrewAI.** The core of the simulation, i.e., the management and operation of the reviewer agent population, was implemented using **CrewAI**[1]. This Python framework is specifically designed to facilitate the creation of autonomous multi-agent systems. Its capability to define:

   - *Agents:* Represent the agents mentioned in the architecture description.
   - *Tasks:* Define the tasks assigned to the agents.
   - *Tools:* `ScrapeWebsiteTool`[2] was used to extract information from the product URL.

   CrewAI allowed abstracting much of the complexity associated with coordinating and executing tasks among multiple agents.
2. **Backend Framework: Flask. Flask** was used to build a lightweight and efficient API, facilitating communication between the user interface, the simulator logic, and the agents.
3. **User Interface: React. React** enabled the development of a dynamic and interactive interface for visualizing the reviews generated by the agents and managing simulations.

[1]https://www.crewai.com/open-source
[2]https://docs.crewai.com/concepts/tools

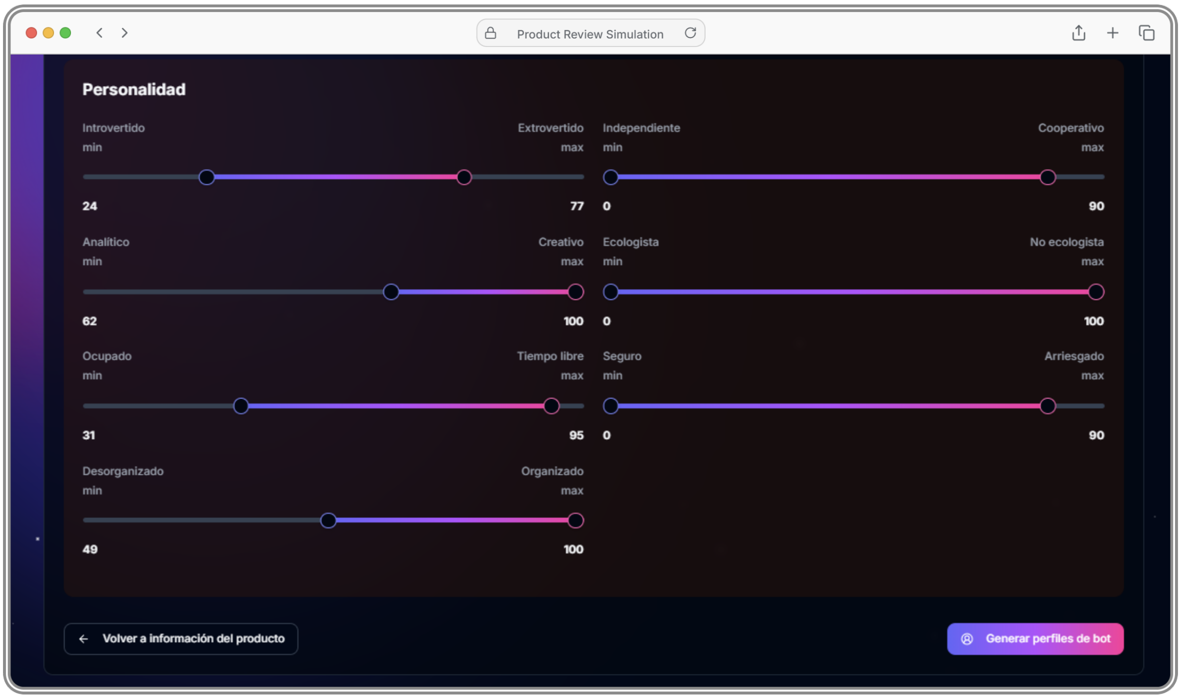

Figure 5: Interface for detailed configuration of agent personality traits.

4. **Large Language Models (LLMs).** The natural language generation capability of the agents is based on integration with large language models. Specifically, **Gemini 2.0 Flash** was used, with its interaction managed through CrewAI.

Together, this technological framework provided the necessary tools to efficiently develop a prototype that credibly simulates the generation of product reviews by configurable agents.

5. Application and Benefits

The main application of this prototype lies in its ability to function as a tool for **preliminary market research and product concept evaluation**. Before making significant investments in development, production, or large-scale marketing campaigns, companies can use this simulation to:

- **Identify Perceived Strengths and Weaknesses:** The simulation reveals which product features are likely to be best valued and which could generate criticism or dissatisfaction in a specific market segment (defined by agent parameters).
- **Test the Reaction of Niche Markets:** By configuring the agent population to represent a particular demographic or psychographic group (e.g., "young university students interested in sustainability," "middle-aged professionals sensitive to price"), it's possible to predict how that specific niche might react to the product.
- **Compare Product Variants:** Simulations could be run for different hypothetical versions of a product (e.g., with different features or prices) to assess which would have a better simulated reception.
- **Refine Marketing Strategies:** Understanding which aspects resonate most (positively or negatively) with the simulated audience can help focus marketing messages on the most relevant strengths and proactively address potential weaknesses.
- **Reduce Risks and Costs:** By obtaining early simulated feedback, more informed decisions can be made, potentially avoiding costly design or market positioning errors.

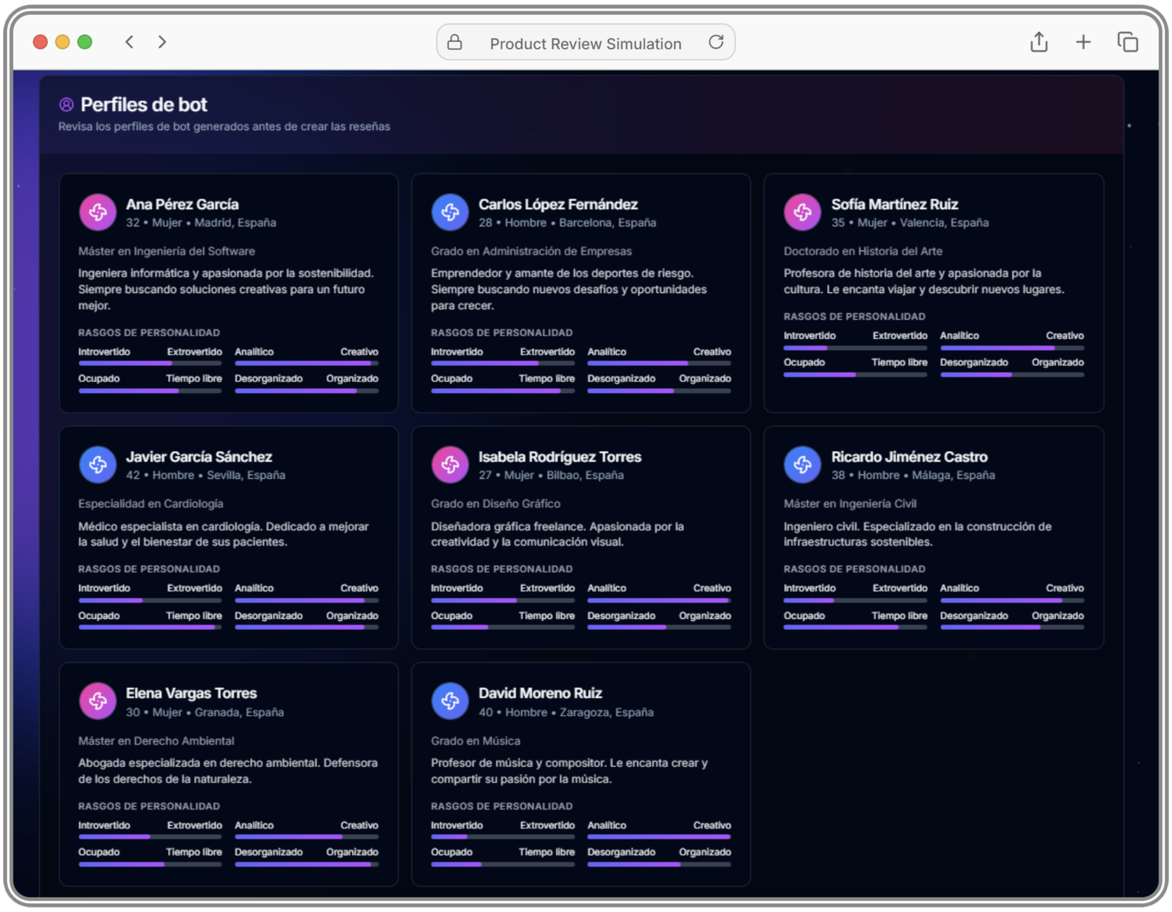

Figure 6: Visualization of the generated profiles and their statistics.

The automated nature of scraping and review generation/analysis allows these insights to be obtained much faster and potentially at a lower cost than traditional market research methods, especially in the initial ideation and design phases. Aspect-based sentiment analysis (ABSA) provides a more granular understanding of opinions on specific product features, which is crucial for informed decision-making [16].

6. Conclusions and Future Work

We have presented a prototype of a multi-agent system for product review simulation. The proposed architecture integrates specialized agents for data extraction, configurable generation of reviews based on simulated consumer profiles, and aggregated analysis of these reviews, culminating in a visual dashboard. This tool offers a novel (To the best of our knowledge there are no similar work in generating artificial opinions in e-commerce platforms) and automated approach to gaining early insight into the possible reception of a product in a target market.

Potential benefits include the early identification of strengths and weaknesses, the ability to test scenarios with different target audiences, and the reduction of risks associated with launching new products.

As future work, we plan to refine the internal models of the reviewer agents, possibly incorporating advanced language models (LLMs) to generate more realistic and nuanced reviews [12]. We will also explore validating the system by comparing simulation results with real reviews of already launched products. Other lines of improvement include expanding agent configuration parameters and

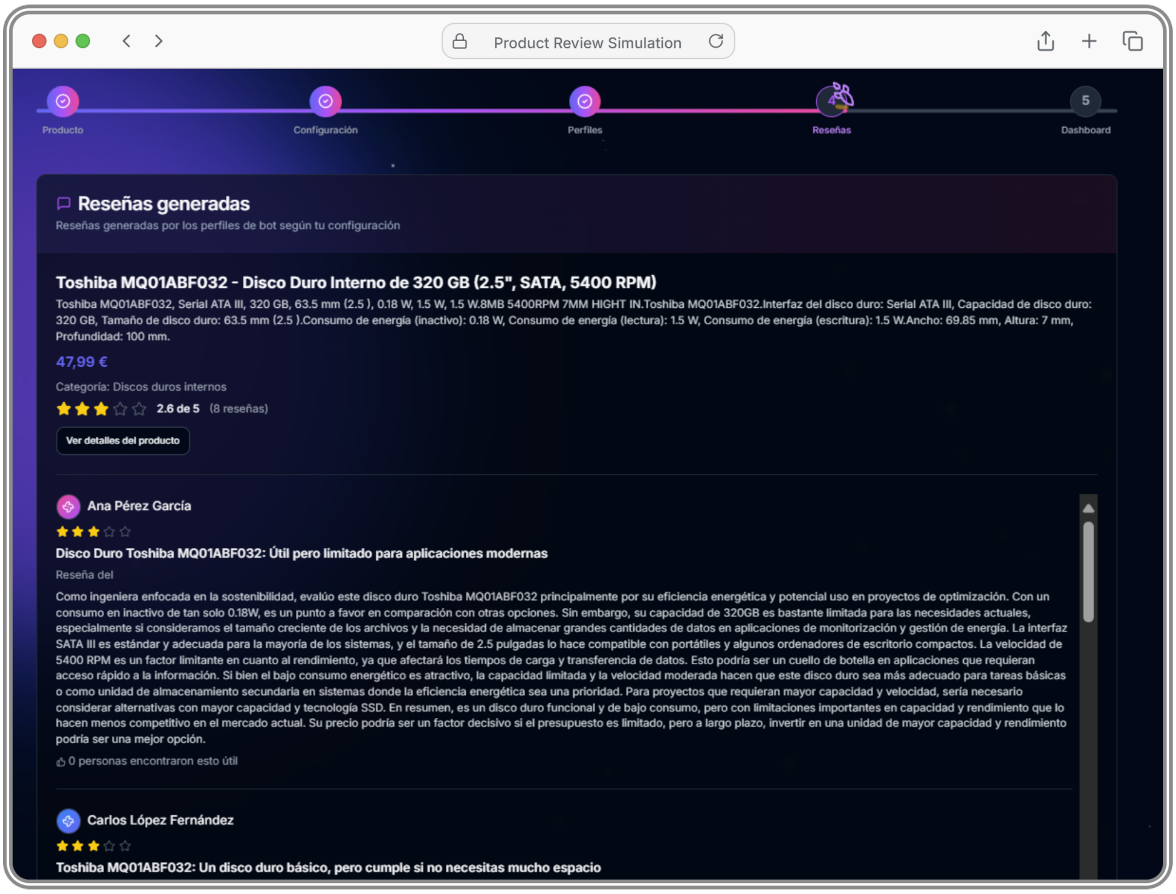

Figure 7: Visualization of the reviews generated by the agent population, presented similarly to online store reviews.

incorporating more sophisticated analyses in the analyst agent (such as detecting emerging trends). Self-improving multi-agent systems, such as the one proposed by [1], offer a promising path to increase the accuracy and utility of our system as it is used with more products and scenarios.

Additionally, a future research line will systematically investigate how agent parameters such as demographics and personality traits concretely influence the prompts provided to the language models and, consequently, the review generation process (Section 3.3).

Regarding the data extraction component (Section 3.1), the robustness of the web scraping process across different website structures remains to be thoroughly evaluated. Future work will include extensive testing and refinement of scraping methods.

Acknowledgments

This work has been partially supported by projects CONSENSO (PID2021-122263OB-C21), MODERATES (TED2021-130145B-I00), SocialTOX (PDC2022-133146-C21) funded by Plan Nacional I+D+i from the Spanish Government, and by the scholarship (FPI-PRE2022-105603) from the Ministry of Science, Innovation and Universities of the Spanish Government. Also, this work has been funded by the Ministerio para la Transformación Digital y de la Función Pública and Plan de Recuperación, Transformación y Resiliencia - Funded by EU – NextGenerationEU within the framework of the project Desarrollo Modelos ALIA.

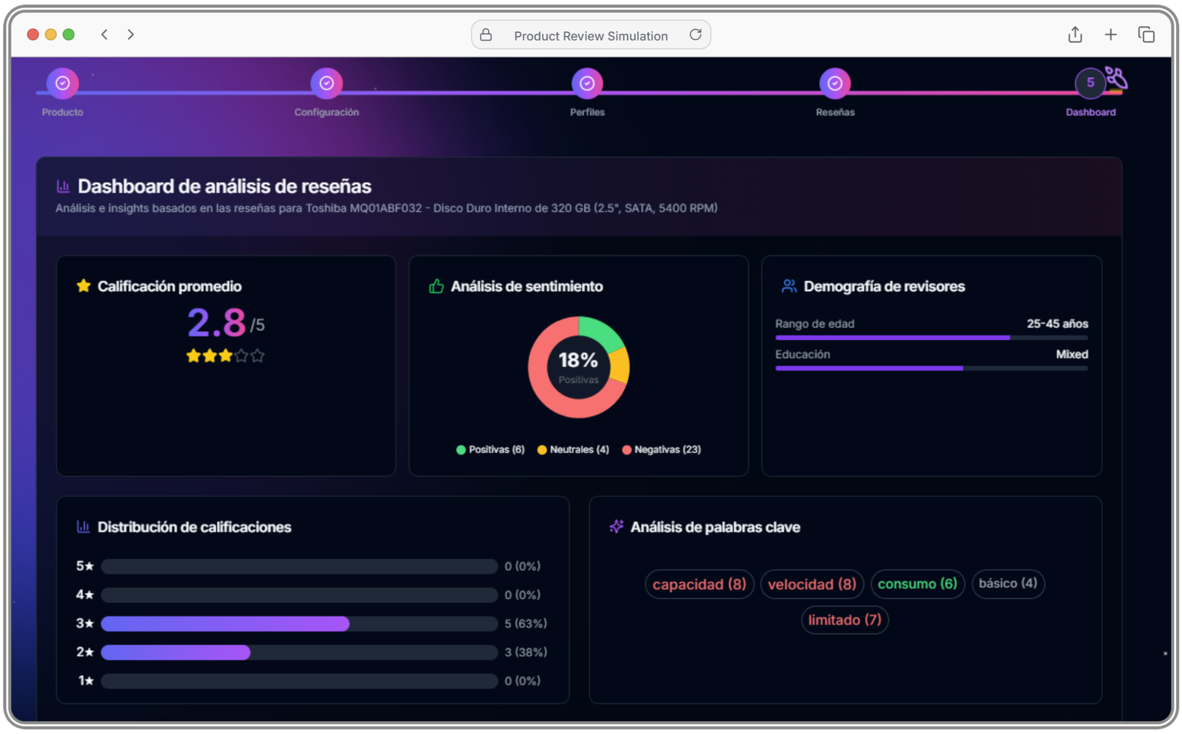

Figure 8: Analysis results presented in a dashboard that gathers all data processed by the Analyst Agent.

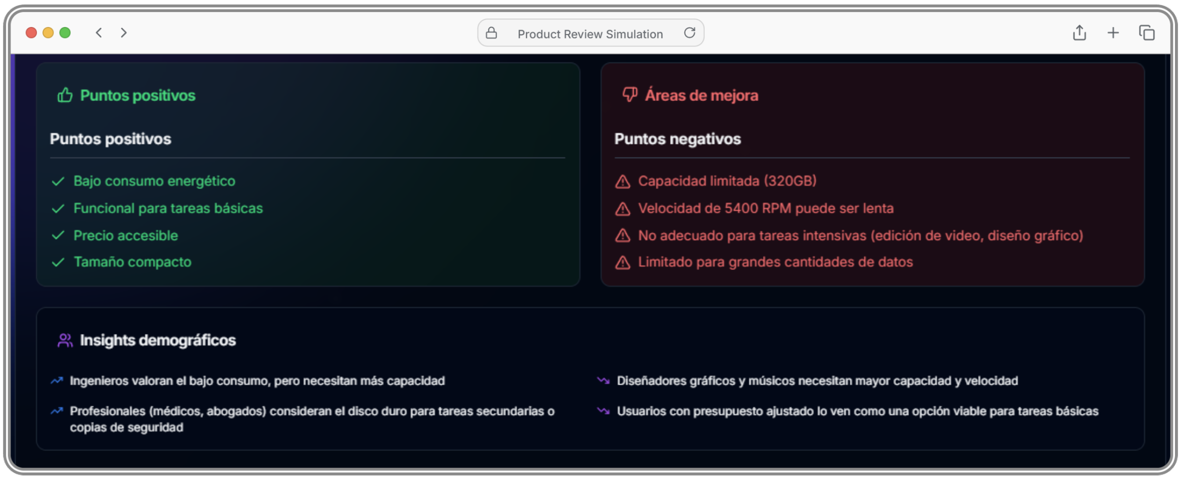

Figure 9: Visualization of the strengths and weaknesses on the dashboard.

Declaration on Generative AI

During the preparation of this work, the authors used GPT-4 and Grammarly in order to: Grammar and spelling check. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] W. Zhao, M. Yuksekgonul, S. Wu, J. Zou, SiriuS: Self-improving Multi-agent Systems via Bootstrapped Reasoning, 2025. URL: https://arxiv.org/abs/2502.04780. arXiv:2502.04780.

[2] T. Guo, X. Chen, Y. Wang, R. Chang, S. Pei, N. V. Chawla, O. Wiest, X. Zhang, Large language model based multi-agents: a survey of progress and challenges, in: Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, IJCAI '24, 2024. URL: https://doi.org/10.24963/ijcai.2024/890. doi:10.24963/ijcai.2024/890.

[3] R. Lin, Y. Gan, T. Lan, X. Liu, Y. Cai, C. Li, D. Luo, Q. Liu, Revisiting aspect sentiment triplet extraction: A span-level approach with enhanced contextual interaction, Expert Systems with Applications 268 (2025) 126126.

[4] Y. Li, B. Zhou, Y. Niu, Y. Zhao, Fine grained sentiment analysis on microblogs based on graph convolution and self attention graph pooling, Applied Intelligence 55 (2025) 1–16.

[5] T. Ge, X. Chan, X. Wang, D. Yu, H. Mi, D. Yu, Scaling Synthetic Data Creation with 1,000,000,000 Personas, 2025. URL: https://arxiv.org/abs/2406.20094. arXiv:2406.20094.

[6] T. Wang, M. Tao, R. Fang, H. Wang, S. Wang, Y. E. Jiang, W. Zhou, AI PERSONA: Towards Life-long Personalization of LLMs, 2024. URL: https://arxiv.org/abs/2412.13103. arXiv:2412.13103.

[7] J. S. Park, J. C. O'Brien, C. J. Cai, M. R. Morris, P. Liang, M. S. Bernstein, Generative Agents: Interactive Simulacra of Human Behavior, 2023. URL: https://arxiv.org/abs/2304.03442. arXiv:2304.03442.

[8] J. S. Park, C. Q. Zou, A. Shaw, B. M. Hill, C. Cai, M. R. Morris, R. Willer, P. Liang, M. S. Bernstein, Generative agent simulations of 1,000 people, 2024. URL: https://arxiv.org/abs/2411.10109. arXiv:2411.10109.

[9] W. Huang, Z. Gu, C. Peng, J. Liang, Z. Li, Y. Xiao, L. Wen, Z. Chen, AutoScraper: A progressive understanding web agent for web scraper generation, in: Y. Al-Onaizan, M. Bansal, Y.-N. Chen (Eds.), Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Miami, Florida, USA, 2024, pp. 2371–2389. URL: https://aclanthology.org/2024.emnlp-main.141/. doi:10.18653/v1/2024.emnlp-main.141.

[10] C.E.Rajaprabha, C. S. Krishna, R.Srikrishnan, B. N. Firthous, S.Nisha, Web scraping based site for product analysis, International Journal For Multidisciplinary Research 6 (2024).

[11] X. Liang, H. Wang, Y. Wang, S. Song, J. Yang, S. Niu, J. Hu, D. Liu, S. Yao, F. Xiong, Z. Li, Controllable text generation for large language models: A survey, 2024. URL: https://arxiv.org/abs/2408.12599. arXiv:2408.12599.

[12] A. Slobodianiuk, S. Semerikov, Advances in neural text generation: A systematic review (2022-2024), 2025.

[13] W. Zhang, Y. Deng, B. Liu, S. J. Pan, L. Bing, Sentiment Analysis in the Era of Large Language Models: A Reality Check, in: Findings of the Association for Computational Linguistics: NAACL 2024, 2024. URL: https://aclanthology.org/2024.findings-naacl.246.pdf.

[14] L. Cabello, U. Akujuobi, It is Simple Sometimes: A Study On Improving Aspect-Based Sentiment Analysis Performance, in: Findings of the Association for Computational Linguistics: ACL 2024, 2024. URL: https://aclanthology.org/2024.findings-acl.394.pdf.

[15] M. Kumar, L. Khan, H.-T. Chang, Evolving techniques in sentiment analysis: a comprehensive review, PeerJ Comput. Sci. 11 (2025) e2592.

[16] Y. C. Hua, P. Denny, J. Wicker, K. Taskova, A systematic review of aspect-based sentiment analysis: domains, methods, and trends, Artificial Intelligence Review 57 (2024). URL: https://link.springer.com/article/10.1007/s10462-024-10906-z.

A. Artificial Opinion Prompt

Prompt used to generate opinions (translated from Spanish):

```
1. Review the following product information:
<JSON Description of the product>.

2. Evaluate the product from the perspective of your personal profile:
<JSON Description of the persona>

3. Generate a review in JSON format that includes:

    • id: a unique number (use <passed index>)
    • bot_id: the ID of the user's profile (<profile ID>)
    • product_id: <product ID>
    • rating: a rating from 1 to 5 stars
    • title: a short and descriptive title
    • content: the detailed content of the review
```