USE OF ARTIFICIAL INTELLIGENCE IN ISTAT TO IMPROVE THE CUSTOMER EXPERIENCE [1]

Gianpiero Bianchi, Carlo Boselli, Mauro Bruno, Alessio Cardacino, Cecilia Colasanti, Mario Magarò, Emiliano Montefiori, Francesco Ortame, Serenella Ravioli, Francesco Rizzo

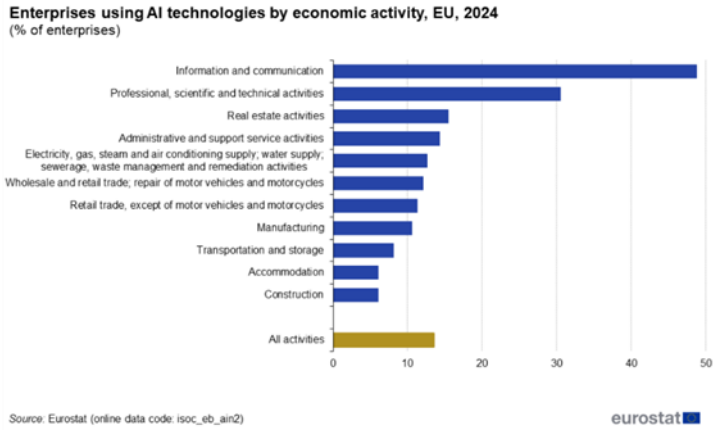
Abstract. The ongoing digital transformation, the evolution of computational capabilities and the widespread use of artificial intelligence (AI) make information management and intelligence strategies increasingly central themes. Starting from the state of the art, the aim of this paper is to explore how these phenomena are interpreted in Istat, through a project dedicated to the use of AI for the data search in Istat's data warehouse, where official statistical data are published. They include, among others, national accounts, household economic conditions, prices, labour and wages, industry, justice, services, health, education, culture, welfare, daily life, public administrations, environment and energy, and agriculture.
The obtained result was to facilitate data searches through the use of an AI-based chatbot. Our goal is to enhance the user experience by assisting users in searching for content already published in the data browser.
Istat, as the official producer of statistics for our country, pays particular attention to the accuracy and quality of the data released. Even if the use of AI is widespread, from a business prespective, it is not yet a mature technology. On the one hand, therefore, the drive for innovation has led us to experiment with this new technology, and Istat is among the first public administrations to offer a similar service. On the other hand, the project, despite being online and validated by the methodological sector, and despite having trained the data only on the Istat domain, has an experimental nature. What we obtained is the first step. We expect to continuously improve our data research by investing in AI.
Finally, we want to emphasize the teamwork that underpinned this project. The IT, communication and methodological departments have worked in a truly synergic way, seeking together a point of balance in the complexity, in the attempt to provide a better public service.

[1] Carlo Boselli and Francesco Rizzo gave their contribution as project manager; Serenella Ravioli was responsible for the dissemination part of the project; Gianpiero Bianchi, Mauro Bruno, Francesco Ortame took care of the methodological point of view of the project; Alessio Cardacino, Mario Magarò, Emiliano Montefiori represented the IT side of the work; Cecilia Colasanti was the IT responsible and reviewed this paper.

1. Introduction

The widespread use of artificial intelligence (AI) impacts both the private and the public sectors. By referring to the "Use of artificial intelligence in enterprises", a recent "Statistics Explained" released by Eurostat[2], shows that AI is developing quickly in many economic activities (Figure 1).

Figure 1 – *Enterprises using AI technologies by economic activity, EU, 2024*.

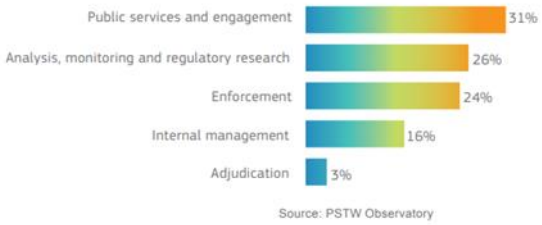


In some economic activities AI is used a lot more than in others. This might indicate that AI is more relevant for certain activities. In 2024, the information and communication sector (with 48.72%) and professional, scientific and technical service activities (with 30.53%) stood out with the highest share of enterprises that used AI.

For the government AI activities in Europe, "AI watch", the European Commission's Joint Research Centre (JRC), studies the evolution of this phenomenon, relevant for digital services and products in the public administration, health, transport, and education. In particular, the Public Sector Technology Watch (PSTW) provides valuable insights about AI governance and AI solutions designed to enhance services within the public sector[3]. As shown in Figure 2, 31% of AI public applications are referred to as "Public services and engagement".

[2] https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Use_of_artificial_intelligence_in_enterprises
[3] https://interoperable-europe.ec.europa.eu/sites/default/files/document/2024-04/PSTW%20-%20AI%20and%20blockchain%20in%20the%20European%20Public%20Sector.pdf

Figure 2 – *AI use cases distribution according to Process and Application Type, EU, 2024*.



Within the public services and engagement domain, service personalisation emerged as the most prevalent application, accounting for 18%. This also marks the highest application type across the entire repository. Engagement management (5%), service integration (6%), data sharing management (1%) and others (1%) represent less prevalent applications.

Naturally, the National Statistical Offices are involved in the use of AI-driven innovation to improve part of the official statistics production.

The use of an AI-based system for the dissemination of official statistics requires a very high level of commitment to validating the model's results. This is especially true when using generative AI components, which can produce not entirely accurate outcomes and must be employed carefully and sparingly. For this reason, there are very few instances of major international institutions that have developed systems or prototypes capable of ensuring adequate levels of reliability. One such example is the project launched by the International Monetary Fund called StatGPT, which is still under development. Similar to the system developed by Istat for IstatData, it uses artificial intelligence models applied to different SDMX data nodes or repositories used by international and national organizations to identify relevant datasets.

Narrowing the field to Istat, several AI-based projects have been started.

This paper aims to explore the AI-IstatData project, which focuses on leveraging artificial intelligence to enhance data search capabilities within Istat's data warehouse, where official statistics are published.

2. The Istat's corporate data warehouse

IstatData is the corporate data warehouse that disseminates the aggregate results of all socio-economic surveys conducted by the Institute. At a strategic level, our data warehouse must serve as the sole dissemination point for official statistical production. Therefore, there is an ongoing a process of consolidation and migration of the different Istat web systems towards IstatData.

In recent years, the data warehouse has been enriched with information, and approximately 3,000 multidimensional tables are currently disseminated, for a total of 2 billion aggregate values. These numbers are expected to grow.

Although the information is organized into themes to make the identification of the tables easier, the growth of topics covered has required the identification of innovative technological solutions for data research, for the benefit of researchers, journalists, institutions and more general users. The AI-IstatData project was therefore born with the aim of improving the research tool and the customer experience in customing the aggregate information available on the Institute's corporate data warehouse through the use of artificial intelligence. Specifically, to enhance semantic search capabilities an AI-based search chatbot was integrated, allowing users to interact with it. Through the same function we have enhanced the use of the user guide. Our goal is to improve the user experience by supporting them in searching for content already published in the data browser. The glossaries and information dictionaries already available on the Istat website will be integrated into the AI system. This enhancement will improve the user interface and chatbot interactions, enabling the AI to also incorporate graphs and maps in response to user queries. Furthermore, in the second phase of the project, it will be possible to maintain the history between the user and the chatbot in order to allow a more effective "dialogue".

The main advantages can be summarized as follows:

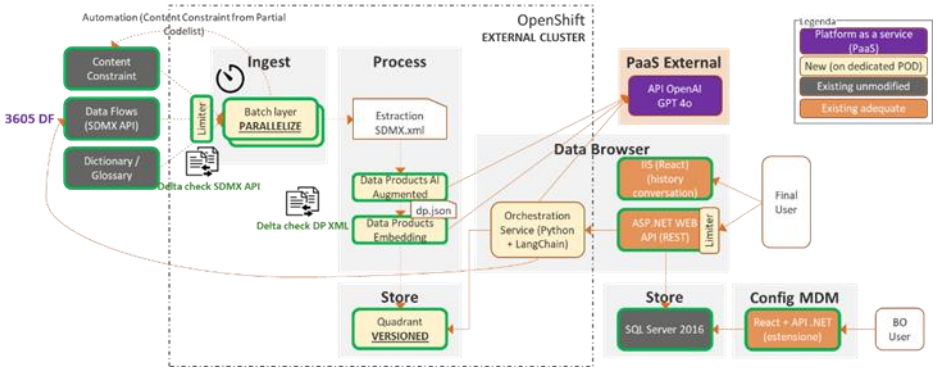
- It allows searches in simple and commonly used words, using natural language (greater ease of use);
- It makes communication smoother through integrated suggestions;
- It explains why data sets are selected by the system and how they are connected to the question (greater transparency). In this way, the user trust can be maintained high.

This chatbot not only improves the search, but also the use of tools like graphic viewers and maps.

2.1. The IT point of view

IstatData is based on the open-source tools "Data Browser" and "Meta & Data Manager" developed by Istat and on the international data transfer standard, recognized by the International Standard Organization, SDMX (Statistical Data and Metadata eXchange) for the exchange and sharing of statistical data and metadata. For those who need to use these data for other developments, the platform also makes machine-to-machine access available via a specific API (Figure 3).

Figure 3 – *IT architecture.*



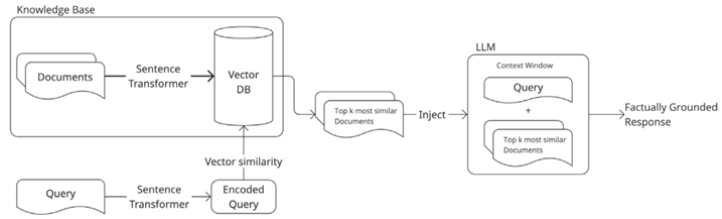
In addition to the use of AI, which represents a great innovation, the IT architecture is also based on cutting-edge technical standard.

2.2. The methodological point of view

Istat, as the official producer of statistics for the country, pays particular attention to the accuracy and quality of the data released. Even if AI is widely used, it is a constantly evolving technology. On the one hand, therefore, the drive for innovation has made us invest in AI (we are among the first PAs to offer a similar service); on the other hand, while the project is now live and has been methodologically validated, using exclusively Istat-sourced data, it retains an experimental character.

System Overview - The chatbot implements a retrieval-augmented generation (RAG) pipeline (Lewis et al., 2020). When a user submits a query, a GPT-4o classifier first assesses whether the query pertains to a dataset search. If not, the system displays a polite clarification or redirection. Otherwise, it proceeds to retrieve candidate datasets.

Figure 4 – *Scheme of a RAG system.*



Standard RAG pipeline.

Figure 4 shows the structure of a standard RAG pipeline. The knowledge base is the set of vectors representing the pool of texts we want to extract. In this case, the knowledge base comprises the complete set of Istat public datasets, each represented by multiple sentence-level dense embedding vectors and a sparse vector representing keywords for exact matching. Texts are transformed into vectors using sentence-level embedding models, specifically, OpenAI's *text-embedding* models. The whole set of vectors is stored in a vector database. Next, a natural language query is embedded using the same embedding model, and its embedding vector is compared to the vectors in the database to extract the most similar measuring vector similarity. This step aims to extract datasets that are semantically similar to the text in the user's query. After the most similar vectors are extracted, they are mapped back to the original datasets, creating a list of candidate datasets. This is list of result displayed in the user interface. This list and the user's query are injected into a lightweight GPT-4o-mini model, which synthesizes a coherent, context-aware, user-friendly response.

Evaluation - There has been a careful evaluation process, particularly towards the relevance of the extracted datasets given user queries, including manual labelling of results with several degrees of relevance.

To test and enhance the search algorithm and its relevance ranking, two distinct types of questions were developed: (1) 340 *specific* questions, each linked to a single correct answer within a strictly relevant dataset, and (2) 516 *generic* questions for which the number of pertinent datasets is broader.

Specific questions were developed by domain experts and data modelers, both familiar with the respective data structures. They formulated these questions by analyzing the contents of specific data structures. Generic questions were developed by domain experts and internal Istat platform users who work closely with end-users and have experience with the types of queries most commonly submitted.

Regarding relevance testing, for the specific questions, it was possible to perform a direct comparison between the search system's results and the predetermined set of correct answers. For the generic questions, however, a manual, point-by-point review of the coherence of the system's top 20 answers for each query was required. This human evaluation was conducted by domain and data warehouse experts, as the number of relevant answers is higher and they can span different thematic areas.

Retrieval Strategy – The retrieval strategy was defined empirically based on the system's performance in the relevance tests (Table 1). Initially, a hybrid—yet non-hierarchical—model combining dense and sparse vector components was used. This approach achieved a 65% objective match rate for specific queries within the top five results. For generic queries, an average of only 57.4% of the top 10 answers were

deemed correct by the domain experts and data warehouse evaluators, with 14% considered non-relevant and a significant 34% being non-relevant and exhibiting clear characteristics of hallucination.

Analysis of the generic responses led to a breakdown into two subsets of comparable size: one derived solely from the dense vector component and the other solely from the sparse vector component. This revealed that the correct answer rate in the first subgroup (dense-only) rose to 81%, while it fell to 17% in the second (sparse-only). Notably, within the sparse-only subset, the rate of hallucinated answers rose to 65%. Furthermore, it was interesting to observe that in the group of answers identified by both the dense and sparse components simultaneously, the share of correct answers rose to 82%, with a noticeable shift of correct results into the top 10 positions.

Given that the answers generated by the sparse component alone were characterized by a substantial rate of hallucinations, the decision was made to adopt a hierarchical model. This model first identifies a preliminary subset of answers using the dense component and then applies a hybrid of both dense and sparse methods within this subset. While other approaches were experimented with, this hierarchical model yielded the most significant increase in correct answers. Specifically, for specific queries, the rate of correct answers in the top 5 results rose to 71% (compared to 65% with the non-hierarchical model). For generic queries, the rate of correct answers within the top 10 results increased to 81%.

Table 1 – *Dataset retrieval accuracy for different scenarios.*

Scenario	Acc. on specific queries (Presence/Top 5)	Acc. on generic queries (Relevant / Top10)
Dense + Sparse non Hier	0.652	0.574
Dense + Sparse Hier	0.713	0.805

Results ranking.

3. A team effort

Artificial intelligence, like any multifaceted phenomenon, requires different professional skills and sensibilities to be governed. Teamwork and deep synergy among different skills become critical success factors especially for a project that uses AI. The introduction of this chatbot is a good example to show the power of team working. Within Istat, the Departments of communication, IT and statistical

methodology have worked in a synergic way, seeking together a point of balance between innovation and rigor. The strategic goal of improving the user experience led us to choose the project manager in the communication area, where the statistical contents of the data warehouse are managed. On the other hand, it was the work of the IT department that allowed the development and release of the project within the expected time and with the required quality, and thanks to the methodological department the rigor of the results was obtained. Naturally, the Department of Legal Affairs also contributed for the risk analysis and for the disclaimer published on the website.

4. Conclusion

The fact that there aren't many analogous experiences makes this project original and pioneering compared to international institutions similar to ours. Istat's experience can be considered food for thought for other producers of official statistics.

First of all, this AI-based project is based on a platform (IstatData) that has centralized and rationalized the statistical data disseminated. Master data management activity is therefore preliminary to any AI project. Indeed, good data quality is an essential element to ensure the quality of the required results.

Connected to the first, the semantic correctness of the data was an element in which Istat invested heavily before applying AI.

Continuous tuning of the results improves the quality of the responses offered to users.

This is a virtuous and meaningful example in which a public institution, putting together its different souls, is providing a better service to the country.

Acknowledgements

The Authors thank the Istat colleagues Mara Cammarrota, Daniela Casale, Cecilia Catalano, Gianluca D'Amato, Catello Di Somma, Gianluca Ferri, Alessandro Ghizzoni, Luigi Malatesta, Fausto Panicali, Manlio Paterniti, Mirko Tassinari, Leonardo Tininini, Michela Troia for their contribution to the implementation of the work.

References

LEWIS P. PEREZ E., PIKTUS A., PETRONI F., KARPUKHIN V., GOYAL N., ... , KIELA D. 2020. Retrieval-augmented generation for knowledge-intensive NLP tasks. In *Proceedings of the 34th International Conference on Neural Information Processing Systems*, Article No. 793, pp. 9459 – 9474.

EUROSTAT. 2025. Use of artificial intelligence in enterprises, *Statistics Explained*.

BRIZUELA A. *et al*. 2024. *Public Sector Tech Watch. Mapping Innovation in the EU Public Services*. Luxemburg: Publications Office of the European Union.

Gianpiero BIANCHI, Istat, gianbia@istat.it
Carlo BOSELLI, Istat, cboselli@istat.it
Mauro BRUNO, Istat, mbruno@istat.it
Alessio CARDACINO, Istat, alcardac@istat.it
Cecilia COLASANTI, Istat, cecolasa@istat.it
Mario MAGARO', Istat, magaro@istat.it
Emiliano MONTEFIORI, Istat, montefiori@istat.it
Francesco ORTAME, Istat, francesco.ortame@istat.it
Serenella RAVIOLI, Istat, serenella.ravioli@istat.it
Francesco RIZZO, Istat, rizzo@istat.it