CARMA 2026

8th Int. Conf. on Advanced Research Methods and Analytics
1-3 July 2026, Valencia, Spain
DOI: https://doi.org/10.4995/CARMA2026.2026.21983

AI for Questionnaires: A Strategic Initiative for the Modernization of Statistical Production[1]

Sabrina Barcherini[1], Guido Drovandi[1], Francesca Rossetti[1]

[1]Department for development of methods and technologies for the production and dissemination of statistical information, Istat - Italian National Institute of Statistics, Italy

How to cite: Barcherini, S.; Drovandi, G.; Rossetti, F. (2026). AI for Questionnaires: A Strategic Initiative for the Modernization of Statistical Production In: 8th International Conference on Advanced Research Methods and Analytics (CARMA 2026). Valencia, 1-3 July 2026. https://doi.org/10.4995/CARMA2026.2026.21983

Abstract

The "AI for Questionnaires" project is a strategic initiative aimed at modernizing the entire lifecycle of statistical survey instruments through the integration of advanced Artificial Intelligence solutions. The primary objective is to generate significant efficiency gains and improvements in the overall quality of collected data by innovating the questionnaire design process. This paper outlines the strategic objectives and the methodological role of the questionnaire as a measurement tool.

Keywords: *questionnaires; data quality; Artificial Intelligence solution; data collection.*

1. Introduction and Project Objectives

The design of a survey questionnaire is a cornerstone of high-quality data production, a principle enshrined in the European Statistics Code of Practice (Eurostat, 2018), which mandates that statistics must accurately and reliably reflect reality. In this context, the "AI for Questionnaires" project represents a profound transformation based on a 360-degree review of all processes linked to the lifecycle of this instrument, with the explicit goal of identifying every area where Artificial Intelligence can generate concrete efficiency gains and significant qualitative improvements.

The modular nature of the project allows for a granular approach to the three main phases of the survey questionnaire: design, technical implementation, and operational field use. Within this framework, AI does not merely intervene to accelerate operations but supports researchers in managing increasing complexities, ensuring high quality standards in the production of statistical data. This contribution is structured into sections detailing strategic objectives,

[1]The article was only possible thanks to the joint work of the authors. Sections 2, 3 and 3.1 have been written by Sabrina Barcherini; sections 3.2, 4 and 5 by Guido Drovandi; sections 1, 3.3 and 6 by Francesca Rossetti.

Article in Press

This work is licensed under a Creative Commons License CC BY-NC-SA 4.0
Editorial Universitat Politècnica de València

methodological analysis, and platform functionalities across the Design, Implementation, and Respondent and Quality macro-areas, and the underlying system architecture based on the "Single Source of Truth" principle.

2. The Questionnaire as a Measurement Instrument

The questionnaire is configured as a sophisticated measurement instrument whose primary objective is the reduction of non-sampling error, particularly response error. Methodologically rigorous design “adapting language, structure, and complexity to respondent characteristics” is essential to minimize measurement error and uniquely communicate the survey's intent (Bradburn, 2004).

This approach is increasingly crucial in today’s survey landscape, characterized by the expansion into new data sources and the adoption of mixed-mode techniques. To govern such complexity, evaluation must focus on the macro-architecture, ensuring thematic sections follow a logical sequence that facilitates the respondent’s cognitive process, and on the micro-structure, ensuring clear and concise language with solid internal consistency between questions, instructions, and definitions.

Furthermore, the definition and control of the navigation flow via skip logic represent a critical dimension; a logical error here is not a simple technical inconsistency but a factor that undermines respondent trust and data integrity. Questionnaire design is thus an iterative process where transparency, consistency, and traceability are critical success factors.

3. Platform Tools and Functionalities

The functional architecture of the project is organized into three macro-areas, each dedicated to presiding over a specific phase of the questionnaire lifecycle through specialized AI tools.

3.1. Design Area

In the Design Area (A1), the system intervenes from the early stages to ensure methodological solidity. Starting from a draft questionnaire, the platform reconstructs and validates the logical structure (A1.1 - Path and logic analysis), identifying respondent profiles and generating real-time flowcharts. This analysis extends to verifying consistency between the questionnaire and all support materials (A1.2 - Cross-document coherence), ensuring that incompatibility plans and instructions are perfectly aligned. To facilitate subsequent analysis, the system automates the creation of the data discharge layout (A1.3 - Record layout definition), standardizing variable names and formats. AI is further employed to examine linguistic registers, identifying semantic ambiguities (A1.4 - Semantic and register audit) and detecting deviations from official classifications (A1.5 - Standard classification mapping). Moreover, the system verifies that the

Article in Press

language and structure remain consistent with the specific data collection technique (A1.6 - Administration technique consistency). To complete this phase, the platform generates technical assets for developers (A1.7 - Implementation asset generation), manages multilingual translations (A1.8 - Translation of questionnaires), and prepares versions adapted for publication (A1.9 - Layout for website version). Finally, a centralized database allows for the mapping of information to identify thematic redundancies across different surveys or administrative sources (A1.10 - Metadata bank and governance).

3.2. Implementation Area

In the Implementation Area (A2), focus shifts to the transition to computer based collection systems. Here, fundamental support is provided by the automatic generation of synthetic respondent profiles: through the creation of minimum and extended sets, AI allows for the exhaustive testing of every logical rule, filter, and text variant before the survey launch (A2.2 - Systematic test profile generation).

3.3. Respondent and Quality Area

In the Respondent and Quality Area (A3), the project presides over the data collection and verification phase. During completion, AI supports respondents and interviewers by integrating search engines for assisted classification (A3.1 - Support for standard classifications) and providing real-time assistance via chatbots (A3.2 - Virtual assistant). Finally, AI supports data validation by generating program code (SAS, R, or SQL) to identify anomalous interviews or suspicious patterns based on known record layouts and deterministic criteria.

4. System Design[2]

The implementation of the project is founded on resource optimization and the robustness of adopted solutions, integrating existing institutional tools like the virtual assistant for the Single Point of Contact and the multilingual translator.

The system is based on a layered and modular logical architecture, designed to ensure operational flexibility and technological independence. The conceptual core lies in the systematization of information: the system transforms unstructured documents into a formalized data model, which serves as a digital "master" and a unique representation of the questionnaire logic. This structure enables integrated functional dependencies, ensuring process consistency as the master model becomes the “Single Source of Truth” for all subsequent automation. For example, the logical decomposition into a directed graph informs both data discharge schemas

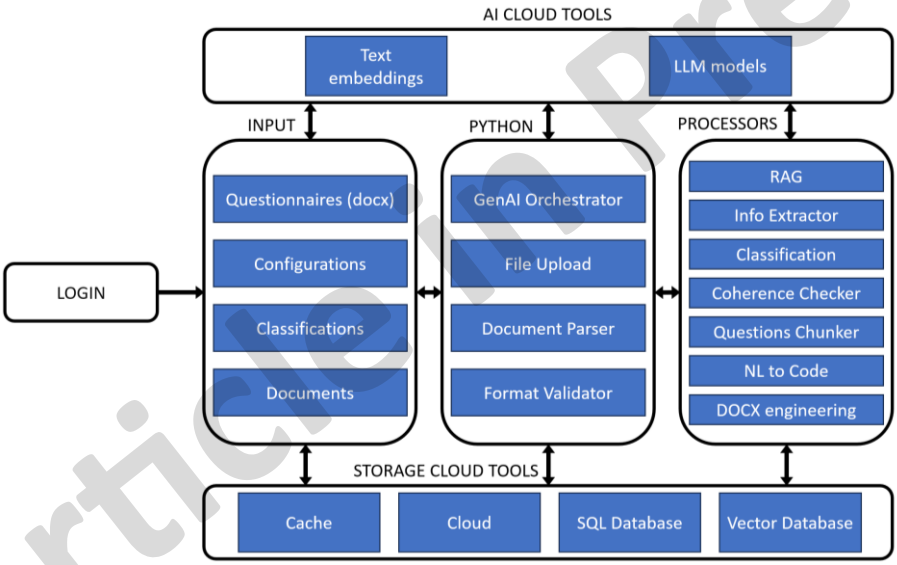
[2] The implementation of the system were outsourced to Deloitte.

Article in Press

and exhaustive testing scenarios. Validation is based on a semantic comparison logic that measures consistency across various materials.

The architecture is designed to be scalable, reliable, and secure, and it will be hosted on cloud services with servers located in Italy to ensure GDPR compliance. The two main components are the Frontend/Backend pillar, developed in Python, and the Processors pillar. The Python frontend/backend is the core and uses the LLM model for validating input documents and extracting metadata. The Processors component, on the other hand, is a container of microservices for advanced AI processing. In this component, the output from the document parser is used to create a hierarchy of all the parts that form a question, check the consistency of the questionnaire, extract the main information from the documents, and improve the automatic generation of content through RAG (Retrieval Augmented Generation).

Furthermore, the architecture is designed to be independent of any specific LLM engine by using the LangChain framework, which allows for greater agility in changing the chosen LLM or GenAI Cloud.



Finally, the approach to quality and governance is metadata-driven, ensuring that AI tools do not have access to raw survey data, thereby strictly respecting data protection regulations (European Commission, 2025).

5. Performance Metrics and Efficiency Evaluation

To ensure the effectiveness and reliability of the AI-driven ecosystem, the project defines a comprehensive set of Key Performance Indicators (KPIs) and specific threshold values. These

Article in Press

metrics are categorized into operational efficiency, methodological accuracy, and structural integrity.

5.1. Operational Efficiency Metrics

The platform is designed to provide high-speed responses and reliable availability to minimize bottlenecks in the statistical production cycle.

Processing and Generation Time: For most automated tasks - including logic analysis, record layout definition, and PDF generation - the processing time for a standard questionnaire is guaranteed to be less than 5 minutes.

Semantic Retrieval Speed: For the centralized metadata bank, semantic search results must be delivered in less than 30 seconds.

Process Optimization (Time Saving): Specifically for automated quality control and program code generation, the system targets a reduction in manual effort of over 70% compared to traditional coding and debugging methods.

5.2. Accuracy and Quality Thresholds

The quality of AI outputs is measured against rigorous methodological standards to ensure the precision of statistical instruments.

Logic Extraction Accuracy: The correct identification of logical entities (filters, skips, and conditions) is set at a guaranteed level of over 95%, with an expected minimum threshold of 90% across complex structures.

Error Detection Performance: The reliability of inconsistency reporting is measured through Precision (minimizing false positives) and Recall (ensuring no critical errors are missed). The target for both metrics is above 90-95%.

Corrective Suggestion Quality: Recommendations for semantic or methodological improvements must be deemed useful and accurate in at least 85% of cases.

5.3. Functional Coverage and Governance

Test Logic Coverage: for automated testing, the synthetic respondent profiles must cover at least 90% of all unique logical paths (edges) within the questionnaire graph.

Data Integrity: The mapping and storage of metadata and informational purposes within the centralized database are subject to a 100% integrity guarantee, ensuring no data loss during the ingestion or mapping processes.

Article in Press

This measurement framework ensures that the integration of AI serves as a robust and transparent support system, elevating the qualitative standards of national statistical production while maintaining strict operational control.

6. Conclusions

Integrating Artificial Intelligence into the questionnaire lifecycle represents an effective response to the increasing complexity of modern surveys.

The project demonstrates how a modular and layered approach allows for the optimization of resources and a reduction in development costs by integrating pre-existing tools. The robustness of the system is guaranteed by the digital "master" which ensures logical consistency from design to final testing. Furthermore, the metadata-based governance model ensures full compliance with privacy regulations and the security of the information heritage.

In conclusion, the system not only accelerates processes but elevates the qualitative standards of statistical production, serving as a model for sustainable and methodologically rigorous innovation.

References

Bradburn, N. M., Sudman, S., & Wansink, B. (2004). Asking Questions: The Definitive Guide to Questionnaire Design — *For Market Research, Political Polls, and Social and Health Questionnaires*. Revised Edition. Jossey-Bass.

Eurostat (2018). European Statistics Code of Practice.

European Commission, "*Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act)"* (C(2025) 5052 final del 29.7.2025), https://digital-strategy.ec.europa.eu/it/library/commission-publishes-guidelines-prohibited-artificial-intelligence-ai-practices-defined-ai-act.

Article in Press