WHAT TOOLS FOR A GOVERNANCE OF AI SOLUTIONS TO SUPPORT OFFICIAL STATISTICS?[1]

Massimo Fedeli, Sara Letardi, Simona Pace

Abstract. Artificial Intelligence (AI) refers to the ability of computers to perform repetitive tasks that would usually require the presence of human intelligence. AI can become a powerful tool in support of Official Statistics, innovating traditional approaches to statistical information production processes throughout the entire data life cycle. To fully exploit the potential of AI initiatives, it is not enough to adopt advanced technology solutions but it becomes strategic to integrate them into a reliable, transparent and consistent framework.
While the speed at which AI solutions are being developed augurs an increasingly rich future of opportunities, it must also find producers of official statistics ready not only to embrace the innovative stimuli but also to know how to guide and govern them to take into account the ethical demands that such a challenge entails. This translates into the need to adequately increase the skills required in the various areas affected by AI while stimulating a reflection on how to equip oneself with a standardized approach and processes to develop AI-based solutions that are in line with the agency's strategic objectives.
Istat has activated several experimental projects for the use of artificial intelligence to support the production of statistical information.
In parallel, through dedicated workshops, training moments and monitoring of the initiatives, it is pursuing a path to build a specific framework for the introduction of AI in the Institute, with the aim of aligning the ethical priorities of AI, ensuring its inclusiveness and sustainability.
In this paper, we present the Institute's experience in the conception, development, and early AI-based use cases, highlighting its challenges and most significant milestones.
We also illustrate how, through the integration of use cases into a reliable and effective monitoring and governance ecosystem, AI applications can be overseen to mitigate their risks without reducing their innovative scope.

1. Introduction

In line with the three-year plan proposed by the Agency for Digital Italy for information technology in public administration (AgID) 2024-2026 (AgID, 2023), the National Institute of Statistics, Istat, has started to adopt Artificial Intelligence

[1] This article is the joint work of the authors, however paragraph 1 is written by Simona Pace, paragraphs 2, 3 and 4.1, by Sara Letardi, paragraph 5 is written by Massimo Fedeli.

(AI) in order to promote innovation in methods and technologies for the production and dissemination of statistical information and the increase of administrative efficiency while respecting the principles of transparency, reliability and sustainability.

In general, public administrations (PAs), with regard to their role in the country's digitization process and their task of protecting the public interest face particular challenges in both the management and proposed regulation of artificial intelligence. Indeed, on the one hand, they play the role of a user, adopting AI technologies to more efficiently perform tasks related to their institutional mission; on the other hand, they directly support the technology, such as through the creation of infrastructure services or joining research programs or through the imposition of procurement criteria (Djefall, 2022).

In its role as a public research organisation, Istat therefore adopts a responsible approach to managing the deployment path of the AI solutions it intends to develop or acquire. AI, if implemented without proper controls and assessments, potentially entails several risks, for example related to the way decisions are processed, privacy violations or security issues.

The main tool to face this new challenge is the creation of a reliable, transparent ecosystem in line with international regulations and best practices, i.e. an AI governance framework that is in line with Istat's own characteristics and values.

In the literature, a framework is defined as a set of tools, rules, processes, procedures and values aimed at ensuring regulatory compliance and alignment with ethical values of the AI solutions to be adopted (Mäntymäki et al., 2022).

Following a path of innovation, the Institute has activated several experimental projects on the use of artificial intelligence to support the production of statistical information. In this context, it is pursuing the construction of a specific governance framework for the introduction of AI in the Institute.

The aim of this article is to outline the path for the construction of the framework, starting with the context in which it is to be developed, defining its scope of action and outlining how it will be activated in the Institute, as well as possible future developments.

The construction of the framework within the Institute will be as "agile" as possible, providing for constant and continuous updates to keep pace with the rapid technological developments in AI and the evolution of national and international regulations.

By integrating planned or developing use cases into a reliable and effective monitoring and governance ecosystem, AI applications can then be overseen to mitigate their risks without stifling their innovative scope.

2. Opportunities and challenges for official statistics

Several definitions of Artificial Intelligence can be found in the literature, in this paper we will refer to the definition contained in the AI ACT, published in the European Official Journal on 12 July 2024, at Art. 3, point 1:

" AI system' means a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments" (European Parliament, Council of the European Union, 2024).

For example, AI makes it possible to analyse and process large amounts of data in real time, automate complex processes, improve operational efficiency and ultimately make more informed and timely decisions. In addition, by assisting researchers in their work, AI can bring innovation to the entire process of statistical data production, from data collection, through respondent support tools such as dedicated chat bots, to data dissemination, such as semantic searches of statistical information or the use of virtual assistants to respond to user queries (HLG-MOS White Paper, 2023). In particular, Large Language Models (LLMs), with their ability to automatically sort textual data into predefined categories, can be used to code and classify collected information under the supervision of researchers, integrating both current methodologies and accumulated experience in organisations. Or they can help to simplify data cleaning and pre-processing activities by identifying and correcting errors, missing values or inconsistencies (HLG-MOS White Paper, 2023).

Finally, as part of the Italian AI strategy, it is planned to promote the definition of approaches for generating synthetic datasets that can be used in specific application contexts. It will be essential to support the use of Privacy Enhancing Technologies (PETs), i.e. digital solutions that allow the collection, processing, analysis and sharing of information while protecting the confidentiality and privacy of the data, thus encouraging participation (AgID, 2024).

At the same time, the adoption of AI raises specific ethical issues (Benanti, 2022)

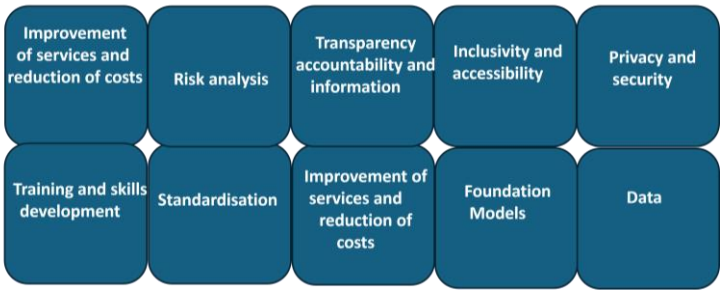
different stakeholders throughout the cycle of an AI solution, be it a product or a service (De Almeida et al., 2021).

In general, trustworthy AI relies on three essential and closely related elements: legality, which ensures consistency with applicable regulations; ethics, which ensures adherence to ethical values and principles; and robustness, which involves careful monitoring and mitigation of the risk of causing intentional or unintentional harm (AI HLEG, 2019).

3. Tools and methods for AI governance.

With the publication of the Artificial Intelligence Regulation (AI Regulation) in the Official Journal of the European Union on 12 July 2024, Europe has set the course to ensure that AI in the European context is trustworthy and respects fundamental rights and the basic principles of democracy and the rule of law. Specifically, the Regulation sets out a number of principles and regulatory solutions that will have an impact in public administration in several areas, from respect for fundamental rights to ethical principles, from methods of legal risk analysis to the need for new skills and organisational models (Fondazione ICSC, 2024).

Figure 1 - *AgID Decalogue, abstract.*



As mentioned in the introduction, currently in Italy, the AgID Three-Year Plan for Information Technology in Public Administration 2024-2026 outlines a general framework for the introduction of AI in public administrations and, in the light of the European regulation, has been supplemented with a decalogue of points of attention[2] for the adoption of AI in PA, a brief summary of which we present in Figure 1.

In addition to the regulatory framework, the development and adoption of AI solutions must necessarily take into account both ethical principles, which are not always promptly translated into legal norms, and the interests of users and, more generally, the requirements expressed by stakeholders.

The set of tools, rules, processes, procedures and values that collectively aim to ensure ongoing alignment with ethical principles, applicable legislation and, more generally, an organisation's own values and goals is called AI governance (Mäntymäki *et al*., 2022). However, although the importance of AI governance has

[2]Cfr. https://www.agid.gov.it/sites/default/files/repository_files/documentazione/Degalogo%20AgID%20IA%20nella%20PA.pdf

been widely emphasised, there are relatively few practical systems for understanding and implementing these principles, i.e. defined AI governance frameworks, including guidelines and best practices for organisations (Mäntymäki *et al*., 2022).

In order to develop a framework at the Institute, we analysed what exists in the international landscape and drew on the experience of other international public institutes and private IT or IT consulting companies.

4. Istat's experience

The definition of the governance of AI applications in the Institute is based on the analysis of the best practices present in the national and international landscape, taking into account the indications of the AgID and integrating the different methodologies used in both the public and private spheres, adapting them to the specific context of a research institution such as the National Institute of Statistics. The aim is to provide practical guidelines and organisational tools for the Institute's units that develop, apply or use artificial intelligence technologies.

From the comparison with the elements described, three pillars have been identified which, taken together, form the backbone of AI governance at Istat:

- Innovation
- Data and rules
- People

Each pillar, as illustrated below, is in turn made up of a series of activities that define and characterise it and which, as a whole, translate the ethical principles and standards that inspire them into lines of action through the identification of specific objectives.

The pillars influence each other through mutual interaction and integration, facilitating communication flows between the various structures involved. For example, regulatory compliance involves the implementation of the "rules" defined in the development of AI solutions throughout their lifecycle. The adoption of new technologies to bring about "innovation" is based on the growth of skills and continuous professional updating, i.e. the training of "people", whether they are developers or users of the applications adopted.

In order to illustrate the main points of the framework to be set up, the following is an illustrative and non-exhaustive summary of the actions identified to implement the AgID indications contained in the Decalogue, shown in Table 1.

4.1. Innovation

The first pillar, Innovation, is characterised by the analysis and evaluation of existing and developing AI applications at the Institute and the definition of tools to ensure that AI solutions are developed, commissioned and monitored in accordance with AgID recommendations and the Institute's values and strategic objectives..

Research and advisory firm Gartner (Ramos *et al.*, 2023), has observed that in the early stages of testing generative AI, organisations focus their attention on technical pilots that deliver incremental improvements, making it difficult to identify and prioritise the potentially most impactful generative AI use cases. Instead, there is a need to manage the governance of requests to implement new AI solutions and identify which use cases are prioritised over others.

In the use case discovery phase, AI solution proposals are identified and catalogued. Then, for each use case, the teams involved, the objectives, the timeframe for implementation, the data used to develop or deploy the AI solution, and any costs to be incurred are identified. It also identifies any risks and/or criticalities that have been or may be encountered in the course of development or deployment, and what mitigation measures can be put in place in relation to the risks and criticalities identified. Finally, the indicators that will be used to measure the effectiveness of the results obtained will be defined.

In summary, based on the analysis of the needs expressed and the innovation plan of the Institute, AI use cases are analysed in terms of defining the objectives, activities, data and technologies used in AI applications.

The path of identification, monitoring and evaluation of the use cases is in line with the indications provided by AgID for the governance of AI, as summarised below:

Table 1 – *Innovation.*

AgID Indication	Requirement	Actions
Improvement of services and cost reduction	Approve new use cases against criteria shared with management	Define and apply priority criteria for use case adoption
Improvement of services and cost reduction	Control and monitor approved use cases	Create and update the Use Case Catalogue
Sustainability	Ensure the sustainable use of resources that are dedicated to AI solutions	Promote the reuse of IA solutions

As an example, here is an extract from the catalogue of use cases used at the Institute for the exploratory phase of approved projects.

Table 2 – *Catalogue of use cases in Istat – extract.*

Use cases	Goals	Benefits
Single point of contact	Implement a generative AI solution within the Single Point of Contact.	Improve the quality of information provided to users in terms of consistency and accuracy. Improve user satisfaction by simplifying access to the system. Improve the Institute's reputation: create a strong identity image, more modern and in tune with the times.
Semantic search engine	Implement a search engine on the new institutional website.	Ability to respond more quickly and accurately to internal and external requests, identification of previously unknown document insights. All this will improve the quality of the service provided to users.
Experimenting with the use of ia to access istat data	Definition of a solution that supports users in the guided navigation of IstatData, through the use of generative AI.	The project aims at enhancing the information assets produced by Istat through the controlled use of a solution based on generative AI. In this project, the main challenge is to produce results that are controlled by avoiding that in the prompting phase the AI algorithms produce uncontrolled effects.

4.2. Data and Rules

The second pillar, Data and Rules, aims to identify and define the processes, procedures and guidelines needed to implement the ethical and regulatory principles of AI governance. Particular attention will be paid to how the data needed for AI applications is collected, manipulated and subsequently disseminated.

In this context, an operational model of AI governance will be defined, identifying roles and responsibilities at both project and organisational levels.

After analysing the recently published AI ACT regulation, the risk level of AI applications will be assessed and possible mitigation measures will be identified. Finally, continuous monitoring of activities and evaluation of results will ensure various aspects, such as alignment with innovation objectives or compliance with regulations, and seize opportunities for continuous improvement, using the various governance tools developed.

For this purpose, it is necessary to continuously collect and update project documentation, in particular for those aspects that concern ethical issues, such as the

data and models used. The use of templates such as scorecards or other *ad hoc* tools makes it possible to monitor the lifecycle of AI solutions and clarify the frame of reference in which they are developed and/or used (Lu *et al*., 2024).

With reference to the AgID indications, several lines of action have been envisaged, by way of example:

Table 3 – *Data and Rules.*

AgID Indication	Requirement	Actions
Risk analysis	Identifying and mitigating risks associated with the use of AI	Defining and constantly updating risk assessment forms.
Transparency, accountability and information	Ensuring compliance with regulations	Recognition and monitoring of project documentation for AI solutions Definition of roles and responsibilities and processes
Privacy and security	Ensure users privacy	Establish training courses on the responsible use of AI

One of the first processes defined in this area concerns the adoption of new use cases, which can be summarised as follows:

- Indication of the strategic objectives or needs of the Institute.
- Creation of a transversal working group
- Experimentation of a Proof of Concept (POC) that includes the minimum functionality and technology required to validate the case.
- Comparison with top management for possible approval and allocation of resources for the development of the POC
- POC implementation
- Evaluation of the results obtained by top management and possible approval of the POC to go into production.
- Go to production

As a first step towards monitoring and managing the growing number of solution projects and the resulting AI systems, a transversal organisational structure has been put in place at the Institute, in line with international best practice. The aim is to monitor and link the initiatives of the various operational structures, ensuring alignment with institutional objectives and working closely with the IT and methodology teams for data management and the integration of AI into applications, with the communication teams to promote cultural growth, and with the legal teams to ensure compliance with the relevant regulations.

4.3. People

The third pillar, People, is dedicated to those within the Institute who use AI applications or are involved in the development of innovative solutions.

A fundamental element in building an AI framework is the design of actions to promote training and cultural growth to develop the necessary skills and awareness regarding the use and development of AI solutions.

On the one hand, training activities will be generalist in nature, to bring non-specialist personnel closer to AI issues and their responsible use, thus aiming at training future AI users. On the other hand, they will be aimed at a smaller and more experienced contingent, with the aim of enhancing and consolidating the skills of those who will be involved in the design and implementation of AI solutions.

At the same time, streams of communication with the outside world will be activated in order to strengthen the positioning of the organisation with its stakeholders and contribute to the cultural growth of the country.

According to AgID, the following are planned in this area:

Table 4 – *People.*

AgID Indication	Requirement	Actions
Training and skills development	Promote AI literacy to all staff	Set up courses, seminars Disseminate information material. Assess existing skills and align them with those required by AI.
Training and skills development	Promote Istat's positioning towards stakeholders. Participate in the cultural growth to promote a conscious use of AI systems	Participation in dissemination events on AI; activation of internal communication lines.

At Istat, more than 20 hours of training were provided to staff, both specialised and non-specialised, on various AI-related topics, both general and more specialised. AI ACT refresher courses and participation in national and international events are planned.

The 15th National Statistics Conference, held on 3-4 July 2024, was dedicated to the topic "Official Statistics in the Age of Artificial Intelligence[3]. In addition, an area dedicated to artificial intelligence has recently been activated on the Institute's intranet with the aim of disseminating information on artificial intelligence, illustrating its potential for the Institute and presenting ongoing use cases. This area

[3] https://www.istat.it/en/event/fifteenth-national-conference-of-statistics/

is a work in progress and will be updated as experiments and training events are added. To complete the picture, links to European and national legislation will be collected, as well as AgID indications, and finally, events and talks dedicated to the topic will be reported.

5. Conclusions and outlook

Several experimental projects will be started at the Institute, including participation in research programmes related to the development of the Italian LLM, through agreements with public bodies, universities, and private entities, in line with the Italian Artificial Intelligence Strategy (AgID, 2024). The framework will therefore be activated in the field and in "agile" mode, in order to seize opportunities for continuous improvement and to adapt dynamically to the different application areas of AI solutions.

The next step is to achieve ISO/IEC 42001 - ARTIFICIAL INTELLIGENCE MANAGEMENT SYSTEM CERTIFICATION, which sets out the requirements for establishing, implementing, maintaining and continually improving an artificial intelligence management system.

AI governance is not just about ensuring one-off regulatory compliance, but must maintain ethical standards over time. As such, it is a system that, by its nature, is constantly evolving to keep fe with technological advances and regulatory developments. It is therefore itself subject to continuous monitoring and evaluation to ensure that it is effective, adaptable and in line with emerging objectives and needs, and that it promotes ethical and responsible digital innovation.

References

AGID 2023. *Piano Triennale per l'Informatica nella Pubblica Amministrazione 2024-2026.*

AGID 2024. *Strategia Italiana per l'Intelligenza Artificiale 2024-2026.*

AI-HLEG HIGH-LEVEL EXPERT GROUP ON AI, 2019. Ethics guidelines for trustworthy AI. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

BENANTI P. 2022. *Human in the loop. Decisioni umane e intelligenze artificiali.* Milano: Mondadori.

DJEFFAL C. 2022. Intelligenza artificiale e pubblica amministrazione in Germania. Riflessioni sui principi orientativi di carattere normativo a partire dall'esperienza

tedesca. *Teoria e Critica della Regolazione Sociale / Theory and Criticism of Social Regulation*, Vol 2, No 23, pp. 179-197

DE ALMEIDA P. G. R., DOS SANTOS C. D., FARIAS J. S. 2021. Artificial Intelligence Regulation: a framework for governance. *Ethics Inf Technol*, Vol. 23, pp. 505–525.

EUROPEAN PARLIAMENT, COUNCIL OF THE EUROPEAN UNION 2024. Regulation (eu) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)

FONDAZIONE ICSC CENTRO NAZIONALE DI RICERCA IN HIGH PERFORMANCE COMPUTING, BIG DATA E QUANTUM COMPUTING - SOCIETAL IMPLICATIONS AND IMPACT RESEARCH GROUP - OSSERVATORIO. 2024. *AI Act: Raccomandazioni per un uso responsabile ed etico dell'Intelligenza Artificiale nella Pubblica Amministrazione*. https://www.supercomputing-icsc.it/wp-content/uploads/2024/05/AI_PA_Raccomandazioni_13maggio2024.pdf

GARTNER 2023. *How to Pilot Generative AI*. https://www.gartner.com/en/doc/797246-how-to-pilot-generative-ai

HLG-MOS White Paper 2023, Large Language Models for Official Statistics. Modernstats

LU Q., ZHU L. , XU X., WHITTLE J., ZOWGHI D., JACQUET A. 2024. Responsible AI Pattern Catalogue: A Collection of Best Practices for AI Governance and Engineering. *ACM Computing Surveys*. Vol. 56, No. 7, pp. 1-35.

MÄNTYMÄKI M., MINKKINEN M., BIRKSTEDT T., VILJANEN M. 2022. Putting AI Ethics into Practice: The Hourglass Model of Organizational AI Governance. arXiv:2206.00335

Massimo FEDELI, Istat, massimo.fedeli@istat.it
Sara LETARDI, Istat, sara.letardi@istat.it
Simona PACE, Istat, simona.pace@istat.it