Ref. Ares(2026)3815259 - 13/04/2026

Co-funded by the European Union

One-Stop-Shop Artificial Intelligence and Machine Learning for Official Statistics

Project 101146355 – AIML4OS

Workpackage 4

ESS AI/ML state-of-play and ecosystem monitoring

Deliverable	4.1 – A comprehensive overview of the current situation of the use of AI/ML in the ESS (Snapshot 1)
Month due	April 2026
Type	Report
Prepared by	Destatis, Unit of Artificial Intelligence and Imputation

Email: Aiml4os-Wp4@destatis.de

Deliverable description

This deliverable presents the results of **Survey 1**, conducted as part of **Work Package 4**. It consists of three parts:

- **The Questionnaire (Survey 1)** stored on GitHub as a PDF which can be accessed here
- **A Short Report**, provided in this document, in which we analysed the results of Survey 1
- **An Extended Report (Snapshot 1)**, compiled in Quarto and hosted on GitHub which can be accessed here

Abstract

This report is being submitted in accordance with deliverable D4.1 within workpackage 4. The aim of this deliverable was to provide a comprehensive overview of the current situation of the use of AI/ML in the ESS.

Short Report

1 Introduction

One-stop-shop for Artificial Intelligence and Machine Learning for Official Statistics Project (AIML4OS) is an ESSnet collaborative project involving 16 countries, started on 1 April 2024 and lasts 4 years.

The overall aim is to develop a 'One Stop Shop' for AI and ML in official statistics; that is, a platform which provides a single-entry point for ESS staff to access a coherent set of capabilities for implementing AI/ML based solutions. To learn more about AIML4OS visit: https://cros.ec.europa.eu/dashboard/aiml4os

Work package 4 (WP 4) focuses on the state of play & ecosystem monitoring and gathers evidence of the current state of AI/ML use in the ESS. The WP4 team developed a short online questionnaire designed to get a high-level overview of the key challenges and successes in this area. This report presents its results in the four areas (Figure 1):

A. **Organizational Aspects** (green) addresses ML architecture, integration, ownership, and responsibilities within the NSI.
B. **Quality Standards** (blue) assesses the existence of quality standards for ML processes.
C. **Human Resources** (purple) focuses on ML skill development strategies and staff competencies.
D. **Current Issues and Challenges** (orange) evaluates organizational and technical limitations affecting ML initiatives.

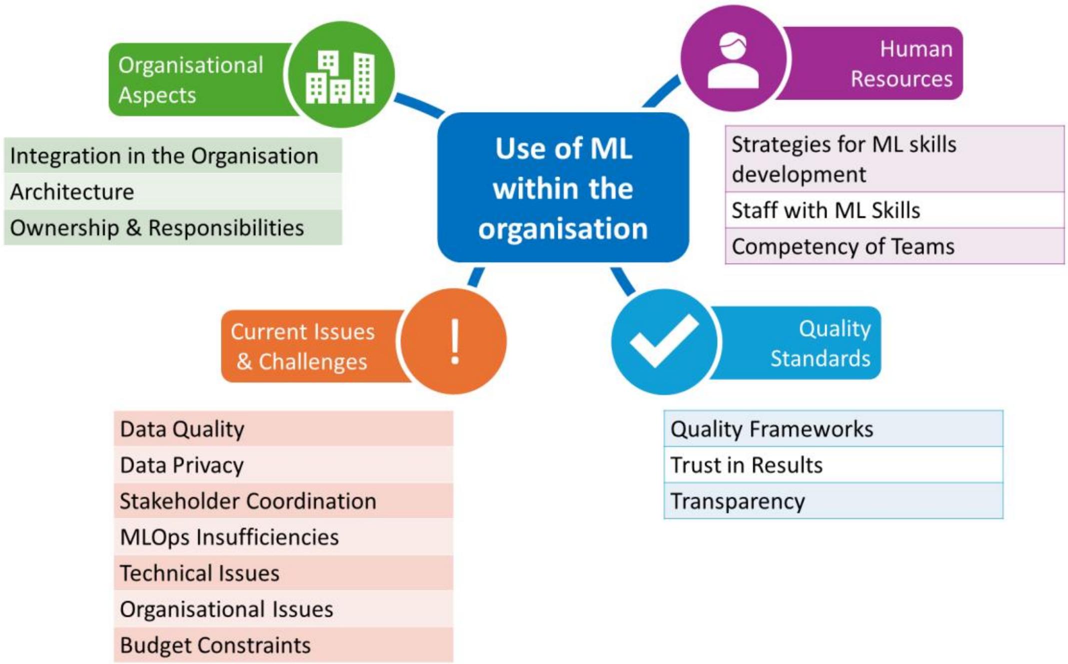

Figure 1 The structure of the survey
Source: own calculations.

2 Profile of Respondents and Institutional Context

This section provides an overview of the institutions and professionals who participated in the survey, offering the necessary context for interpreting the more detailed findings that follow. Rather than presenting responses question by question, the aim here is to summarise the essential characteristics of the participating National Statistical Institutes (NSIs) and their current relationship with Machine Learning (ML).

2.1 Participation and Institutional Landscape

A total of 29 institutions from 28 countries[1] participated in the survey, representing a wide cross-section of the European Statistical System (ESS) as well as a few non-EU members. This broad geographical coverage allows for the exploration of regional patterns and shared challenges in the development and use of ML across different statistical systems.

[1] List of participating countries: Austria, Belgium (2 NSIs), Canada, Croatia, Cyprus, Czech Republic, Denmark, Estonia, France, Germany, Hungary, Iceland, Ireland, Italy, Liechtenstein, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Serbia, Slovenia, Spain, Sweden, Switzerland, and Turkey.

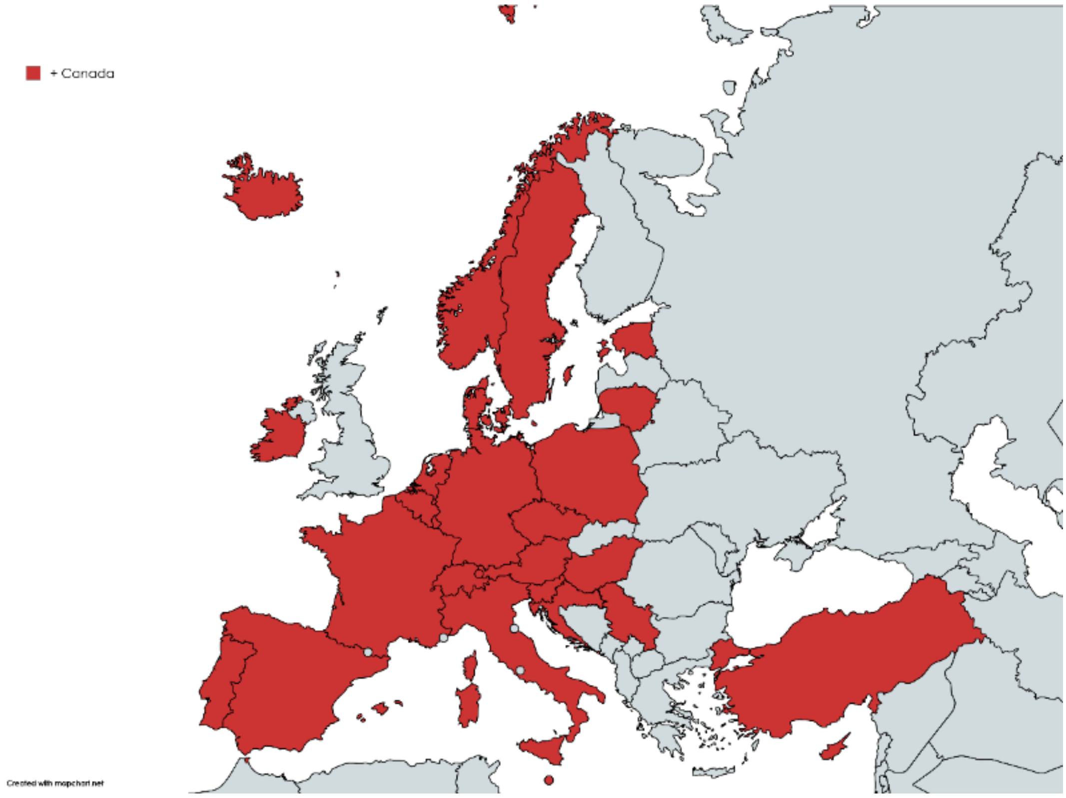

Figure 2 The map of participating countries
Source: own calculations.

2.2 Professional Background of Respondents

The survey was designed to capture perspectives at a strategic level and was therefore intended to be answered by individuals in managerial or decision-making roles. However, it was recognised in advance that responses might also be provided by participants in other roles. For this reason, the role of each respondent (question p3) was collected and is considered relevant contextual information for the interpretation of the results, allowing for a more nuanced analysis of the findings.

The majority of responses came from Managers/Policy Makers (52%) and Statisticians/Data Scientists (41%), reflecting the perspectives of those who either shape institutional strategy or work directly with analytical and methodological development. These two groups together represent almost the entirety of the respondent pool and therefore strongly influence the interpretation of the survey results.

Only around 7% of responses came from Software Engineers or IT Specialists, corresponding to just two individuals. This number is too small to be considered representative of technical or infrastructure-focused roles across NSIs. The survey was intentionally targeted at managerial profiles. Although responses from non-managerial roles were not actively sought, a comparable number of Statisticians or Data Scientists also responded (12 versus 15 managers). As these profiles were not part of the intended target population, their representation is interpreted with caution.

Which option best describes your role in your organization?

p3	Count	Percentage
Statistician/Data Scientist	12	41.38
Analyst/Subject matter specialist	0	0.00
Manager/Policy Maker	15	51.72
Software Engineer/ Information Technology Specialist	2	6.90

Figure 3
Source: own calculations.

2.3 Current Use of Machine Learning

ML adoption is already widespread: around 90% of participating institutions reported using ML in some capacity.

Only three institutions are not currently applying ML techniques, which positions ML as an emerging but increasingly common component of statistical production and research across NSIs.

2.4 Reasons for Not Yet Using ML

Among the small number of institutions not currently using ML, the reasons reflect practical constraints and strategic caution rather than resistance. The explanations fall into five themes:

- **Resource constraints**, especially in smaller offices.
- **Exploratory or pilot-only activity**, without production deployment.
- **Lack of immediate need**, where traditional methods remain fully adequate.
- **Concerns about interpretability and transparency**, essential in official statistics.
- **Skills gaps**, although capacity-building efforts are underway.

All institutions expressed openness to future adoption, underscoring a shared recognition of ML's potential.

2.5 Understanding and Definition of ML

Institutions already using ML were asked whether they agreed with a proposed definition[2]. The results show:

[2] "Machine Learning often (but not always!) involves:
- non-parametric statistical methods,
- recognizing patterns and relationships,
- a focus – at least in so-called supervised learning – on prediction (and less on explanation),
- methods which are used to answer specific questions without being explicitly given the solution,
- thus are generally data-driven rather than model-driven
- and are characterized by the fact that the solution space (meaning the hypothesis space) is often so large that it (approximately) contains all patterns and relationships"

- A **majority agreement** (about 58%), indicating that the definition resonates with prevailing practice.
- **No explicit disagreement**, suggesting no fundamental conceptual opposition.
- A **high level of neutrality** (about 35%), which likely reflects the absence of a formal internal definition of ML in many NSIs.
- A small number (8%) who considered the definition too narrow.

Overall, this points to a community where ML is widely used but where conceptual frameworks are still developing, leaving room for future standardisation or shared guidance.

3 Organizational Aspects

3.1 Evolution and Integration of ML

ML adoption is often driven by practical needs.[3]

Respondents were presented with a series of statements to choose from which describe potential approaches for the adoption of machine learning. The statement most chosen by respondents was "In response to specific business challenges" (18 NSIs). This was followed by "Gradually, through bottom-up initiatives" (17 NSIs), indicating the importance of bottom-up innovation in this space. A large number of respondents also chose the statement "By identifying and implementing high-impact use cases" (15 NSIs). The statement, which was chosen with the lowest frequency, was "Through a dedicated digital strategy" (just 3 NSIs).

Overall Conclusion: ML adoption is mainly problem-driven and bottom-up, with strategic digital roadmaps still the exception rather than the rule.

3.2 Organizational Architecture Supporting ML

Organizations tend to support ML through experimental and flexible structures, rather than fully embedded system-wide architectures (Figure 4).

- **Innovation hubs and pilots lead.** "Innovation Hub and Pilot Projects" was the most frequently perceived architecture, showing that experimentation remains the dominant mode of engagement.
- **Strategic and collaborative models are also visible**. "Centralized, Strategic Approaches" and "Demand-Driven, Collaborative Approaches" were also widely selected, suggesting that many institutions combine top-down structure with bottom-up responsiveness.
- **Hybrid expert-driven models are present but moderate**. They appear, but are not yet dominant, indicating partial but not full integration of expertise.
- **Data-driven decision-making is least cited.** Surprisingly, "Data-Driven Decision-Making" received the lowest number of selections, suggesting that ML is still not central to organizational governance or operational choices.

Overall Conclusion: Most NSIs rely on pilot environments and mixed strategies, showing early-stage ML maturity and limited integration into core decision-making processes.

[3] Multiple answers were allowed in this question.

This result derives from a multiple-response question: respondents could select more than one option, meaning totals exceed the number of participating NSIs. Rather than representing mutually exclusive categories, the responses provide an indication of the relative priorities and focus areas across NSIs.

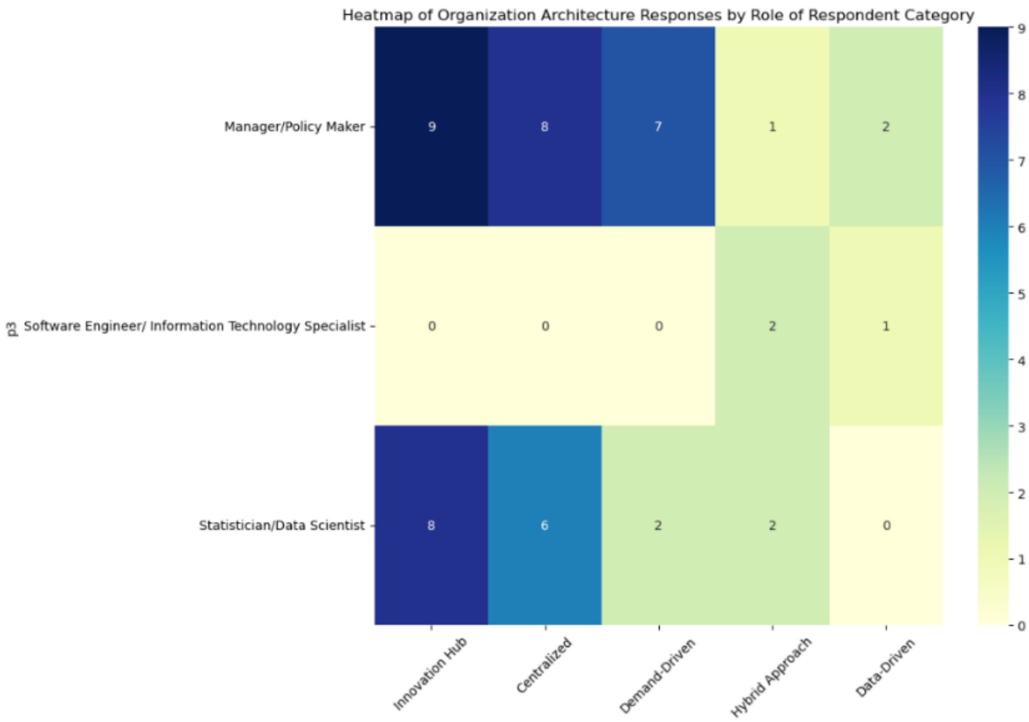

Figure 4
Note: Heatmap showing the distribution of respondent counts across different organizational architecture responses, with respondents listed vertically and architectural responses categorized horizontally.
Source: own calculations.

4 Quality Standards for Machine Learning

Across the NSIs surveyed, the development of ML-specific quality standards is still at an early stage. Only a minority of institutions report having dedicated regulations, guidelines, or frameworks for ML, while the majority operate without formalized structures (Figure 5). This reveals a significant maturity gap, especially considering the sensitive nature of official statistical production and the importance of transparency, accountability, and reproducibility in ML-based processes.

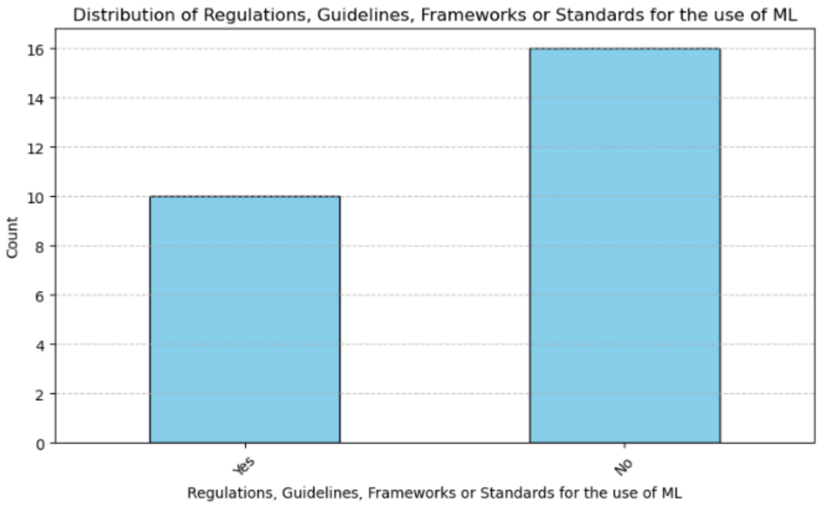

Figure 5
Source: own calculations.
Among the organizations that do have ML-specific standards, three major themes emerge.

1. **Compliance and data protection are the most common focus areas.** Several institutions explicitly mention GDPR, privacy rules, restrictions on non-compliant generative AI tools, and ethical or information-security safeguards. These measures ensure that ML activities respect existing data protection obligations and mitigate risks associated with sensitive data use.
2. **Methodological and quality frameworks play a central role.** Some institutions apply general quality frameworks or map ML processes to existing structures such as the GSBPM, while others provide methodological guidelines for model inputs, outputs, and documentation. Support from methodology or quality teams also appears as a mechanism to reinforce rigor and traceability.
3. **Governance structures are beginning to take shape.** A smaller number of NSIs have defined ML-specific roles, established AI working groups to review use cases, or created collaborative arrangements between data science teams and business areas. These governance practices indicate a gradual institutionalization of responsibilities, although they remain far from universal.

For those without existing ML standards (16 NSIs), three NSIs plans to introduce them in the foreseeable future. Thirteen institutions are currently discussing the creation of guidelines, pipelines, or good practices, such as developing standards for automatic coding, defining rules for the use of cloud-based large language models, or establishing broader governance frameworks for ML. Only a very small minority (3 NSIs) reported no plans to introduce any standards, suggesting a general movement toward more structured and responsible ML integration across the ESS.

Communication of standards varies across institutions.[4] Internal newsletters, communication channels (6 NSIs), and training materials (5 NSIs) are the most frequently used methods, supplemented by events or workshops (4 NSIs). Managers tend to rely more on formal documentation and reports, while statisticians more often mention the use of NSI websites. Overall, dissemination practices are still evolving, even in institutions that already have standards in place.

Regarding how standards were introduced (question b6), most institutions describe a mixed process: both day-to-day operational needs (bottom-up) and management initiatives (top-down) played a role

[4] Multiple answers were allowed in this question.

(Figure 6). A smaller share cites predominantly bottom-up development driven by practical challenges, while top-down-only approaches are the least common. This suggests that ML governance emerges most naturally when it is anchored both in strategic direction and concrete methodological needs.

Was the introduction of standards/frameworks a bottom-up process (inspired by the needs from daily activities) or a top-down process (management's instruction to issue a guideline, for example)?

b6	Count	Percentage
Bottom-up process	3	30.0
Top-down process	2	20.0
Both	5	50.0

Figure 6
Source: own calculations.

Overall conclusion: The landscape of ML quality standards in official statistics is marked by early development, strong emphasis on data protection, modest progress in methodological formalization, and emerging—though still uneven—governance structures. The clear intention of many institutions to adopt such standards in the near future indicates that quality assurance for ML is rapidly becoming a priority across the ESS.

5 Human Resources and Skills for Machine Learning

The results of this section reveal a workforce that is strong in traditional statistical expertise but faces significant challenges in the technical and operational competencies required for modern Machine Learning (ML). Across institutions, respondents consistently point to a substantial gap between well-established statistical foundations and the more engineering-oriented skills needed to build, deploy, and maintain ML systems.

The survey indicates that statistical and mathematical foundations are by far the strongest area (Figure 7). Respondents widely recognize descriptive statistics, hypothesis testing, imputation, and related competencies as part of their institutional core. Domain expertise within statistical business areas is also solid, reflecting long-standing traditions in official statistics. However, the way respondents answered suggests that some may have reported their own knowledge rather than true team-wide institutional competence—especially visible among IT specialists, who do not acknowledge core statistical concepts despite their ubiquity in NSIs. This points to a potential communication gap between professional groups.

In contrast, data engineering, data management, programming, and modern ML techniques show uneven maturity. Basic data cleaning and integration are common, reflecting continuity from traditional processes, but more advanced skills—such as pipeline development, big data handling, cloud platforms, or feature engineering—are much less established. The picture is similarly mixed in data visualization: institutions feel comfortable with dashboards and general-purpose tools, but specialized visualizations for ML remain limited.

The most significant weaknesses appear in Machine Learning algorithms, techniques, and MLOps. Respondents frequently report lacking sufficient skills in key areas such as supervised and unsupervised learning, model evaluation, ensemble methods, or dimensionality reduction. MLOps stands out as the single weakest domain, with extremely few institutions able to integrate models into production environments, monitor them, or manage deployments through automated workflows. This is a critical bottleneck: without MLOps capacity, even well-developed ML prototypes cannot be reliably scaled.

Organizational skills—cross-functional collaboration, knowledge sharing, strategy alignment—show a distinctive divergence between managerial and technical staff. Managers tend to emphasize team structure and cross-unit cooperation, while statisticians identify knowledge sharing as the most visible competency but do not perceive strong alignment around ML strategy or workflows. These discrepancies suggest a perception gap: managers see emerging structures, but operational staff do not yet experience them as fully implemented.

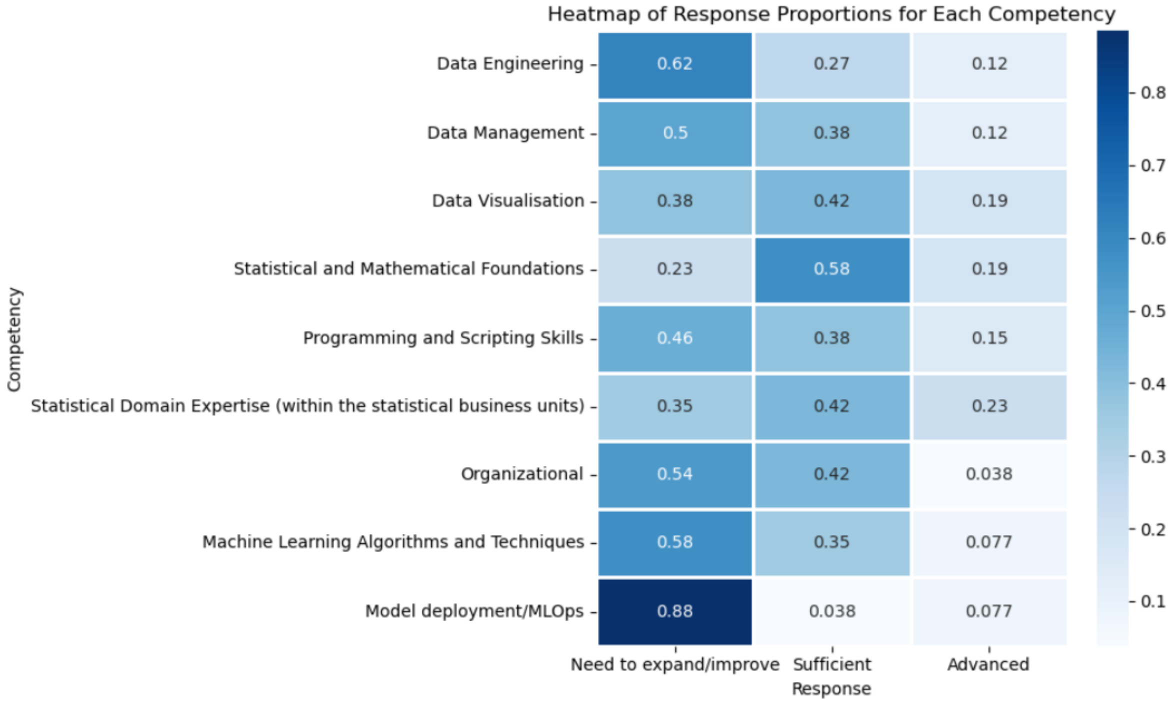

Figure 7 Note: Heatmap displaying the response proportions for each competency. The values within each cell represent the share of respondents, with the totals across each row summing to 1.
Source: own calculations.

Institution size adds important context (Figure 8). Most responding agencies fall in the mid-range but differ widely in the number of data scientists they employ. A small number of large institutions maintain sizable ML teams, while many medium or smaller NSIs operate with very small groups or, in one case, none at all. This variation reinforces the unevenness of ML maturity across the ESS.

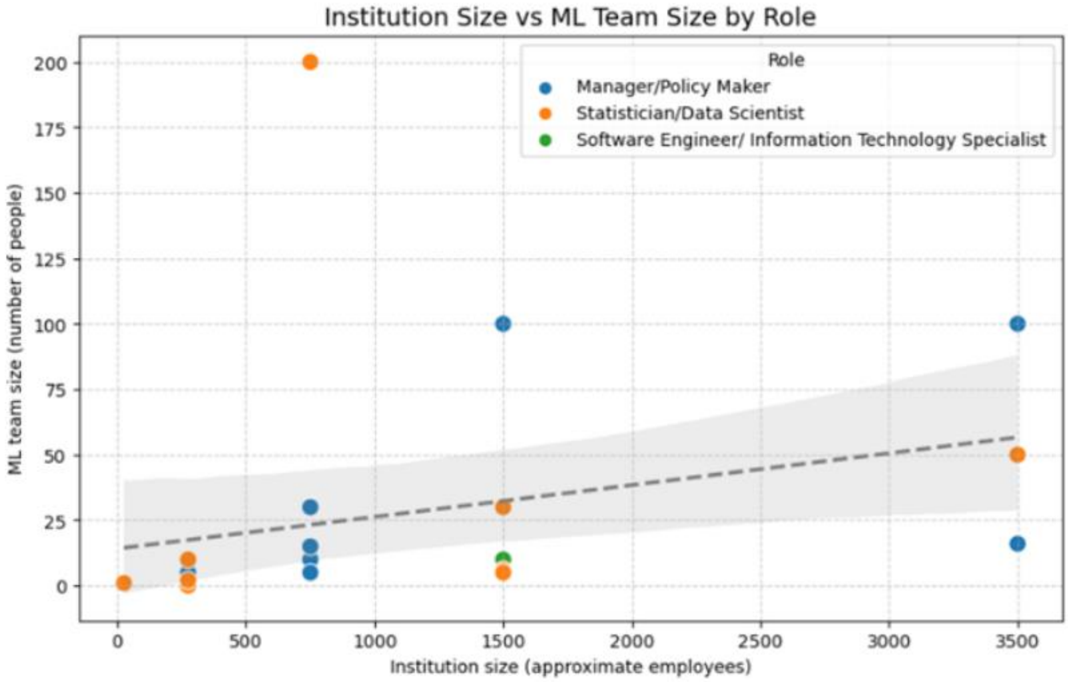

Figure 8
Source: own calculations.

Despite these gaps, there is strong intention to expand ML-related teams. Most institutions (23 NSIs) indicate a clear ambition to grow their ML workforce, reflecting increasing recognition of the strategic importance of these competencies.

Regarding how institutions build ML skills[5], collaboration with universities and research institutions emerges as the most effective strategy (17 NSIs), widely endorsed across all roles. Recruitment of new data scientists follows as a complementary approach (16 NSIs). Training strategies vary by profile: managers favour internal training, while statisticians and engineers value external courses and online learning. More informal mechanisms—such as hackathons, conferences, or communities of practice—receive a mixed response, suggesting they are not yet fully embedded in organizational culture.

Overall conclusion: NSIs possess a strong traditional statistical foundation but face considerable challenges in the engineering, computational, and operational competences that ML requires. The most acute gaps lie in ML techniques and especially MLOps, which are essential for scalable and trustworthy implementation. Institutions are aware of these weaknesses and are actively seeking to expand teams, invest in partnerships, and strengthen training efforts. The results point to a workforce in transition—anchored in classical statistical strengths but still building the modern capabilities needed for ML-enabled statistical production.

6 Current Issues and Challenges

This section examines how organizations perceive the main structural, technical and cultural issues that limit their ability to use Machine Learning, and how these perceptions relate to the level of investment directed at addressing them.

[5] Multiple answers were allowed in this question.

6.1 Organizational Limitations to ML Adoption and Corresponding Investment Efforts

The organizational analysis is based on eight dimensions: internal stakeholders, internal resistance, external resistance, ownership and responsibilities, lack of strategy, lack of coordination, lack of frameworks, and budget constraints (Figure 9).

The findings show that organizations generally acknowledge the potential of ML but face substantial barriers rooted in strategic and structural shortcomings. Financial constraints and the absence of a clear ML strategy emerge as the most significant limitations, frequently reported as moderate to severe. Coordination challenges, unclear ownership structures, and the lack of formal frameworks also appear as recurring problems, suggesting that many institutions are still developing the internal infrastructure necessary to support sustained ML adoption. In contrast, cultural and behavioural issues, such as internal or external resistance, tend to be perceived as relatively minor and do not constitute a major obstacle for most respondents.

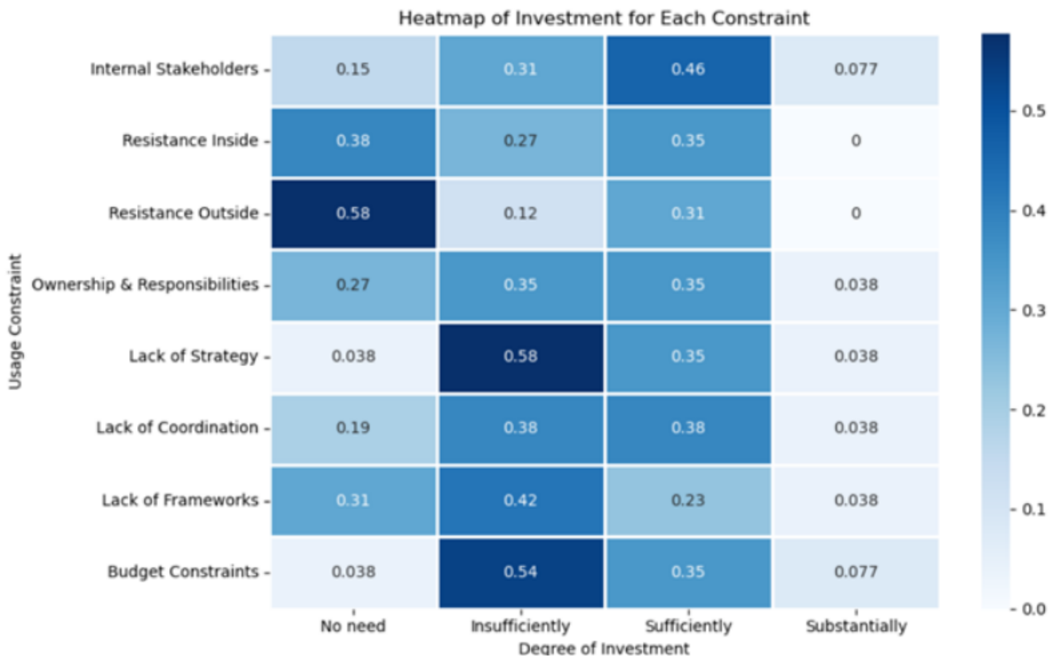

Figure 9
Note: Heatmap displaying the proportion of respondents of usage constraints across degree of investment. The values within each cell represent the share of respondents, with the totals across each row summing to 1.
Source: own calculations.

When comparing these limitations with the reported investment patterns, a consistent misalignment becomes apparent. Although strategy is identified as one of the most critical weaknesses, most organizations acknowledge that their investment in this area remains insufficient. Similar gaps are visible in coordination, frameworks, and ownership structures, where the recognition of organizational shortcomings is not yet matched by adequate resources or reform efforts. Budget constraints stand out as an especially paradoxical challenge: they are widely viewed as the strongest limitation, yet they are also the area where insufficient investment is most frequently reported, reinforcing a cycle in which financial scarcity limits the capacity to overcome the very constraints it creates.

Cultural dimensions paint a different picture. Since internal resistance, external resistance, and stakeholder engagement are broadly perceived as minor issues, investment levels in these areas largely reflect the low level of concern. Many organizations report that no additional investment is required, and those that do invest generally consider their efforts adequate.

Overall conclusions: the relationship between perceived limitations and corresponding investments suggests that organizations are moving towards ML adoption but remain hampered by foundational institutional gaps. Awareness of structural constraints is well established, yet investment patterns indicate that many institutions have not fully committed the resources necessary to resolve them. As a result, ML adoption remains caught between intention and capability: organizations are willing and conceptually prepared, but their strategic and financial preparedness has not yet reached the maturity required for systematic implementation.

6.2 Technical Barriers to Machine Learning Adoption and Corresponding Investment Efforts

The technical dimensions assessed were: availability of ML algorithm skills, programming skills, staff to deploy and maintain ML systems, data fragmentation, maturity of MLOps practices and tools, access to computational hardware, access to broader hardware and software, data quality and bias, availability of suitable training data, and data privacy and protection concerns (Figure 10).

The findings reveal that the most critical technical constraints relate to human capital and operational readiness. Many organisations report that limited in-house expertise in ML algorithms and programming constitutes a major barrier. Despite widespread awareness and growing interest in ML, practical hands-on experience remains uneven, resulting in a reliance on external consultants or experimentation that rarely scales into production. These skill gaps extend beyond development into the operational phase: a shortage of staff capable of deploying, maintaining, and integrating models into business systems emerges as one of the most severe and persistent obstacles. This suggests that while model prototyping is increasingly common, the step towards reliable, sustainable production environments remains underdeveloped.

Closely connected to these challenges is the immaturity of MLOps practices. Many organisations report that inadequate tooling for deployment, versioning, monitoring, and lifecycle management severely limits or even prevents effective ML use. The absence of mature operational pipelines often leads to stalled projects, with promising prototypes unable to transition into stable, reproducible solutions. Data fragmentation also remains a recurring problem: although often rated as a moderate rather than severe limitation, the presence of multiple silos and inconsistent data formats makes integration and model development unnecessarily complex. Infrastructure-related issues show a more heterogeneous pattern. While some organisations have adequate computational hardware or cloud access, others still struggle with outdated or insufficient resources, which slows training and experimentation. More general access to hardware and software appears less problematic overall, although full integration with ML-specific tools remains uneven.

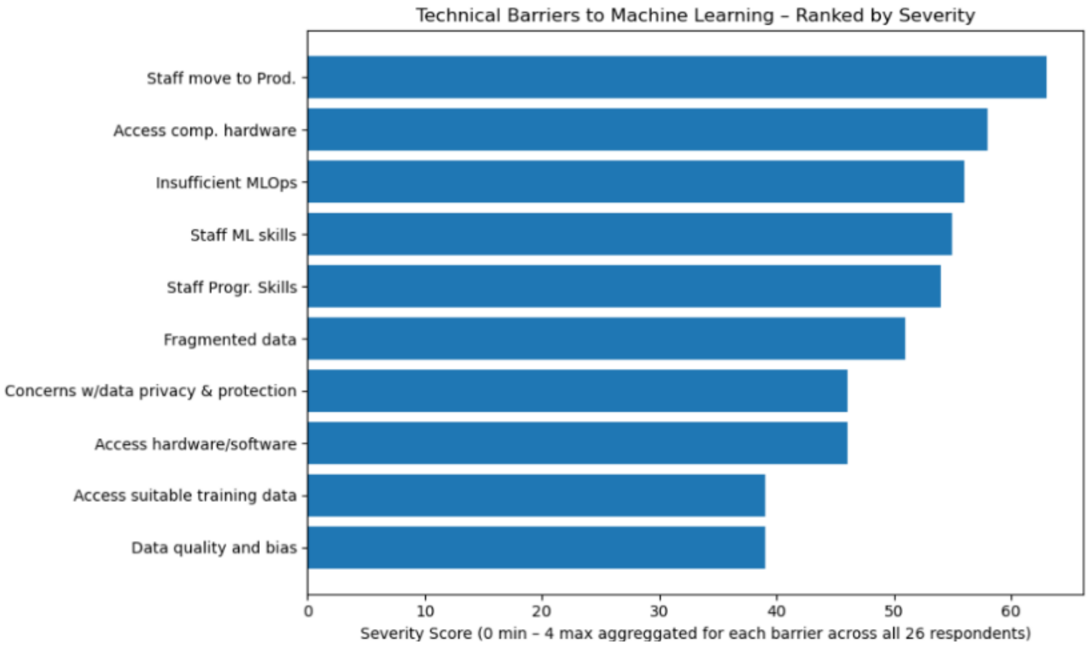

Figure 10
Source: own calculations.

Data challenges also persist but tend to be experienced as moderate rather than prohibitive. Concerns about data quality and bias are widespread, reflecting growing awareness of their impact on model performance and fairness. Access to labelled or relevant training data remains an obstacle for many, indicating that the availability of data is often insufficiently aligned with the specific needs of ML projects. Data privacy and protection represent a more nuanced limitation: while they can restrict data availability and experimentation, many organisations have already developed strong compliance frameworks that help mitigate their operational impact.

When these technical limitations are compared with the corresponding levels of investment, a clear pattern emerges. Across the areas of ML skills, programming skills, deployment capacity, and MLOps maturity, most organisations report that current investment levels are insufficient. Even though these dimensions are consistently identified as the most severe technical bottlenecks, they are not yet receiving the sustained, structured investment required for meaningful progress. This gap suggests that while organisations understand where the main technical obstacles lie, they may lack long-term strategies or resources to address them systematically.

Investment in improving data integration is similarly uneven. Some institutions have begun efforts to address data fragmentation, but the overall scale remains limited relative to the importance of having harmonised, interoperable data systems. Infrastructure investments present a more balanced picture, with several organisations having already secured adequate hardware and software for typical workloads. However, advanced or large-scale ML workloads remain out of reach for many, indicating that infrastructure maturity varies considerably across the landscape.

In the areas of data quality, bias mitigation, and access to training data, investment patterns are mixed. A minority of organisations report mature data governance practices, while others continue to struggle with inconsistent or insufficient datasets. This diversity likely reflects differences in organisational maturity and analytical use cases. By contrast, investment in data privacy and protection is generally assessed as sufficient, reflecting the influence of regulatory frameworks such as the GDPR and the institutionalisation of compliance-oriented data management practices.

Overall conclusions: the results show that organisations have not yet aligned their investment strategies with the technical areas of greatest strategic importance. The most significant limitations—skills availability, deployment readiness, and MLOps capacity—are precisely the areas where investment remains insufficient. Meanwhile, comparatively mature areas such as privacy and general infrastructure continue to receive steady support. This imbalance suggests that many organisations are still focused on foundational IT and compliance requirements, while the more advanced operational components necessary for robust ML adoption are only beginning to be developed. Reaching higher levels of ML maturity will require a deliberate shift towards targeted upskilling, operational pipelines, and integrated data ecosystems, enabling organisations to move beyond experimentation and fully leverage ML in the production of trusted statistical insights.

Supporting information

This site is linked to the workpackage page on CROS: WEBSITE LINK