BANCA D'ITALIA

EUROSISTEMA

Questioni di Economia e Finanza

(Occasional Papers)

Applying artificial intelligence to support regulatory reporting management: the experience at Banca d'Italia

by Canio Benedetto, Sara Crestini, Alessandro de Gregorio, Marco de Leonardis, Andrea del Monaco, Daniele Gulino, Paolo Massaro, Francesca Monacelli and Lorenzo Rubeo

April 2025

Number 927

BANCA D'ITALIA
EUROSISTEMA

Questioni di Economia e Finanza

(Occasional Papers)

Applying artificial intelligence to support regulatory reporting management: the experience at Banca d'Italia

by Canio Benedetto, Sara Crestini, Alessandro de Gregorio, Marco de Leonardis, Andrea del Monaco, Daniele Gulino, Paolo Massaro, Francesca Monacelli and Lorenzo Rubeo

Number 927 – April 2025

The series Occasional Papers *presents studies and documents on issues pertaining to the institutional tasks of the Bank of Italy and the Eurosystem. The* Occasional Papers *appear alongside the* Working Papers *series which are specifically aimed at providing original contributions to economic research.*

The Occasional Papers *include studies conducted within the Bank of Italy, sometimes in cooperation with the Eurosystem or other institutions. The views expressed in the studies are those of the authors and do not involve the responsibility of the institutions to which they belong.*

The series is available online at www.bancaditalia.it .

ISSN 1972-6643 (online)

Designed by the Printing and Publishing Division of the Bank of Italy

APPLYING ARTIFICIAL INTELLIGENCE TO SUPPORT REGULATORY REPORTING MANAGEMENT: THE EXPERIENCE AT BANCA D'ITALIA

by Canio Benedetto*, Sara Crestini*, Alessandro de Gregorio*,
Marco de Leonardis*, Andrea del Monaco*, Daniele Gulino*,
Paolo Massaro*, Francesca Monacelli* and Lorenzo Rubeo*

Abstract

This work describes the approach taken by the statistical function within Banca d'Italia, which manages regulatory reporting data, in using artificial intelligence/machine learning (AI/ML) solutions to support the data management process. It reviews the nine studies carried out so far (six of which have already been implemented in our day-to-day operational processes) to improve statistical processes in three areas: data validation, data enrichment, and process efficiency and automation. For each work, we analyse the business case and the goals, illustrate the solutions identified, and discuss the results. On the basis of the experience gained so far, we draw the main lessons learnt with regard to methodological, organizational, reputational, and procedural implications. Finally, we outline the most promising directions for future research and the implementation of new solutions.

JEL Classification: C32, C38, C45, C81, G21, M15, M53, O33, Y40.

Keywords: regulatory reporting, banking reporting, data quality, data management efficiency, data enrichment, information management, statistical production, artificial intelligence, machine learning.

DOI: 10.32057/0.QEF.2025.927

* Bank of Italy, Statistical Data Collection and Processing Directorate.

1. INTRODUCTION[1]

Central banks base their analysis and policy decisions on a large array of data that they either collect from financial and credit institutions (in brief, "reporting agents" – RAs) based on their regulatory power (data on balance sheets, payments services, interest rates, asset management, prudential and resolution indicators, etc.) or acquire from national and international authorities, commercial providers or statistical institutes through data sharing agreements or contracts (non-regulatory information).

While the availability of non-regulatory information mainly implies a procedural, organizational, and sometimes financial effort on the central banks' side, the acquisition of regulatory data also requires RAs to establish and maintain complex and costly workflows. This is why it is the responsibility of the central banks' statistical functions not only to increase the efficiency of their internal data management procedures but also to reduce the reporting burden by streamlining the data quality process to limit the cross-check activity of RAs and relying on alternative sources to fill in data gaps, instead of resorting to their regulatory power to order the submission of missing data, whenever possible.

In light of the above, different studies have been carried out by Banca d'Italia (and others are under way) with the aim of using Artificial Intelligence/Machine Learning (AI/ML) solutions to limit the areas of manual intervention in data management processes, to increase the accuracy of the quality checks by reducing the number of "false positives"[6], and to collect new financial information from alternative sources. It is worth remarking that most of the studies have already been implemented in the day-to-day operational processes and help reduce the (human) effort required by such processes and increase the quality of the data.

On the basis of the experience accumulated so far, the adoption of AI/ML solutions to central banking statistical processes is not straightforward and presents a number of challenges. They include the application of ethical principles when designing the model[7], the need to ensure model transparency to be accountable to RAs when (potential) reporting errors are detected, collaboration with domain experts to correctly interpret outputs, and the difficulties in recruiting specifically skilled staff able to deal with highly sophisticated methods, that can only be partially mitigated through intensive training. Additionally, the success of an approach based on AI/ML heavily depends on the quality and completeness of the data fed into the model during training. For this purpose, in addition to the application of appropriate data cleansing methodologies, robust data governance[8] and sound management practices are also of paramount importance.

The aim of this paper is to share our experience within the statistical function of Banca d'Italia in terms of exploring the potential of AI/ML solutions for the management of regulatory reporting and of ancillary data repositories (e.g. entities and securities registers). Besides touching upon the methodological aspects

[1] We wish to thank Laura Graziani Palmieri, Laura Mellone and Roberto Sabbatini (Banca d'Italia) for their valuable inputs on a previous draft of the paper. The views and the conclusions are solely those of the authors and do not necessarily reflect those of Banca d'Italia.

[6] This refers to cases in which a data error is incorrectly detected.

[7] For example, it is important to avoid any form of RA "discrimination" and to take energy consumption in due consideration when developing the AI/ML solution.

[8] Regarding statistical information used for institutional purposes, the Statistics Committee of Banca d'Italia is the high-level body in charge of the broad coordination of the information needs (in terms of content and quality level) and the design of sound methodological solutions with a view to a multi-purpose and user-friendly use of the data. The Committee also discusses how to innovate the statistical production process in order to improve its efficiency and effectiveness. Further in M. Casa, L. Graziani Palmieri, L. Mellone e F. Monacelli "The integrated approach adopted by the Bank of Italy in the collection and production of credit and financial data" Questioni di Economia e Finanza (Occasional Papers), 667 (2022.)

of the solutions found, we will focus on the rationale behind the choices made and reflect on the constraints or critical issues that have emerged. In particular, we will:

- describe the approach adopted so far to develop AI/ML solutions for the management of regulatory reporting (Chapter 2);

- take stock of the research done so far with regard to three main areas of activity: data validation (Chapter 3), enrichment of register data (Chapter 4), and data management efficiency and automation (Chapter 5). In these three chapters, we summarise the results of papers that have already been published by highlighting the business problems and goals, illustrating the solutions identified, and discussing the results;

- analyse the lessons learnt from the point of view of the statistical function (Chapter 6);

- outline the most promising directions for future research and further implementation of AI/ML solutions (Chapter 7).

2. THE APPROACH OF BANCA D'ITALIA IN THE USE OF AI/ML SOLUTIONS FOR DATA MANAGEMENT

Central banks regularly collect quantitative credit and financial data from RAs according to a set of predefined reporting requirements that are not static but are continuously re-evaluated in relation to the changing economic conditions and new risks that might emerge because of specific shocks. For example, the 2008-09 global financial crisis drew the attention of the policy-maker to securities portfolios and detailed information on loans and related guarantees. Similarly, the impact of extreme climate events highlighted the importance of closing the information gaps on climate change indicators. Although reviews of statistical information requirements are periodically conducted, new data needs often build on existing surveys, ultimately contributing to further increase the huge amounts of records that are regularly processed by central banks statisticians.

It is important to remark that the new requirements take more and more the form of granular data, i.e. data characterised by a very high level of detail thanks to the many attributes that are defined. Indeed this is a request put forward more and more often by central banks' analysts because, among other things[9], granular data allow to (1) more effectively identify the fundamental causes of economic phenomena by providing a drill-down of the statistics according to the defined aggregation variables, and (2) reduce the time-to-market in delivering new indicators, since new insights may be obtained "just" as different elaborations of the same elementary reporting. Dealing with granular datasets implies much larger volumes of information, new types of diagnostic (quality) checks, adequate investments in IT platform and new statistical skills.

When granularity corresponds to the level of "individual information"[10], it also requires complete and high-quality "reference data"[11] providing the necessary aggregation variables. That is why, along with the collection of individual information, registers (notably, entities and securities registers) play a prominent

[9] Further considerations concerning granular data can be found in J. Brault, M. Haghighi and B. Tissot *"Granular data: new horizons and challenges", IFC Bulletin No 61 (*July 2024). https://www.bis.org/ifc/publ/ifcb61.pdf

[10] By "individual information" we refer to information presented in its most elementary form, as defined in the specific context. Aggregated data is the result of the sum of different pieces of individual information.

[11] By "reference data" we refer to data used to classify or categorise other data in order to provide context. For the purpose of this work with "reference data" we also refer to "master data" (i.e. data describing the business entities). DAMA-DMBOK: Data Management Body of Knowledge (2nd ed.). Data Management Association_2017.ISBN_978-1634622349

role in the statistical data processing field within central banks. Registers are based on different sources. Usually, RAs are requested to provide the information at their disposal on customers and instruments (i.e. the reported units). However, for the reference data that reflect intrinsic characteristics of the "unit", and therefore do not vary with respect to the RA, in order to reduce the reporting burden, central banks can make an effort to exploit other external sources (e.g. Business Registers or financial information providers), or even to fill the informative gaps with their own estimates.

In a nutshell, the availability of reliable and complete data (reference and non-reference) plays a crucial role in central banks' analysis and has an implication for the robustness and soundness of their policy actions and decisions. In this light, a fundamental goal of a central bank's statistical function is to establish a robust statistical landscape characterized by high-quality, complete data as well as efficient operational processes that minimise the reporting burden on banks and other financial intermediaries. Of course, Banca d'Italia is not an exception. Traditionally, it has always placed significant importance on the implementation of a robust Data Quality Management (DQM) process relying as much as possible on fast and automated procedures. The DQM process is based on a multi-layer system of standardised checks designed to validate the accuracy, completeness, and consistency of the information received from RAs. The system comprises formal, deterministic and plausibility checks, each addressing different aspects of data quality: formal checks ensure adherence to data formats and metadata standards; deterministic checks verify the internal consistency of reports based on predefined rules; plausibility checks assess data against historical trends and statistical thresholds to identify outliers. Although some checking algorithms can be tailored to the specific characteristics of each RA or depend on the magnitude of the reported phenomenon, most rules apply to all of them[12]. However, such an approach has been increasingly challenged by the aforementioned growing complexity, variety and volume of the financial datasets. The heterogeneity in business operations among different financial institutions of course makes standardised validation procedures less able to effectively detect nuanced or subtle anomalies; in this context, "one size fits all" is no longer the best performing approach to DQM.

When registers come in the picture (we refer to Banca d'Italia's Entity Register[13] or Securities Register)[14] to complement the "individual" data collection, it is fundamental that also the aggregation variables, and related measures therein, do not show gaps that would otherwise jeopardize the ability to produce fully representative statistics. Filling the gaps that may be encountered in reference data is an activity that the area of the statistical function at Banca d'Italia, which is responsible for such data carries out for the benefit of both the internal data users and the RAs with the goal, respectively, to provide users with complete and accurate information and limit the RAs' reporting obligations.

The trend towards more and more granular data (regulatory) surveys implies the collection of a volume of data that necessarily requires innovative approaches - and the related supporting IT solutions - to effectively scan and process the reported data to ensure a high level of quality while maintaining a high level of operational efficiency in terms of the number of human resources devoted to this activity. Pursuing the efficiency of the operational processes, including the ones concerning data collection and production, is of paramount importance also in terms of the responsible use of financial resources on the side of a central bank for fulfilling its mandate. It is in the tradition of the statistical function of Banca d'Italia to invest in the direction of automating repetitive, massive, time-consuming and error-prone processes related to the data management (e.g. most of the data checks, data base alignment, versioning

[12] More on the management of regulatory reporting at Banca d'Italia in "The integrated approach adopted by the Bank of Italy in the collection and production of credit and financial data" M. Casa, L. Graziani Palmieri, L. Mellone e F. Monacelli (2022) and in "PUMA cooperation between the Bank of Italy and the intermediaries for the production of statistical, supervisory and resolution reporting" M. Casa, M. Carnevali, S. Giacinti and R. Sabatini (2022).

[13] Home/Statistics/Credit and financial reporting/ Entities reference data.

[14] Home/Statistics/services/securities database.

tracking, periodic compilation and dissemination). However, in the new statistical context described above, the volume and variety of the data to be processed requires important investments in new methods, skills and IT solutions.

These goals were included in the 2017-19 the Bank's Strategic Plan[15]; since then, the statistical function of Banca d'Italia has been investigating the potential of AI/ML technologies in enhancing the quality of credit and financial data as well as in improving the efficiency of the operational statistical processes. It is important to remark that when it comes to introducing cutting-edge technologies and methodologies into the operational processes, there must be a careful cost-benefit assessment and any change must be carefully analysed and implemented if necessary. Obviously, as it is already quite established within the AI/ML community and confirmed by the AI Act, it is also important to ensure that it is not AI/ML who "takes" the decisions, rather its role should always be to support a human-based analysis and decision-making process (*human-in-the-loop*).

Furthermore, although the prospect of significantly boosting the efficiency and effectiveness of its statistical processes through AI/ML solutions is undeniably very attractive, the statistical function of Banca d'Italia has avoided an approach driven by uncritical enthusiasm on the potential of the most innovative methodological approaches. Rather, an approach based on carrying out empirical research and the related proofs of concept of the new methods in various areas of the regulatory reporting, carefully planning their actual implementation in the operational processes, has been followed. Indeed, the implementation in the internal statistical processes of AI/ML methods is proceeding following a step-wise approach: firstly, experiments are carried out on individual tasks (proofs of concept); secondly, the costs and benefits of the innovative solutions are evaluated, including the implications on the required skills and IT resources; third, when the benefits overcome the costs and the various risks are analysed and mitigated if necessary, the solutions are actually implemented. Notably, of the nine works summarised in this paper, five[16] have already been implemented in operational processed (although one had to be updated due to changes in the source data, as discussed in par. 6.2) and the implementation of a sixth is already planned[17]. In all cases, we have observed a boost in the efficiency and effectiveness of the day-to-day operational processes.

In brief, the path we are following within the statistical function concerning the adoption of AI/ML solutions to data management can be summarised as: start small, get results, invest in skills and IT resources, then scale up. More specifically, a number of years ago we started with an initial call for interest within the Statistical Data Collection and Processing Directorate aiming at pooling motivated staff to devote their time to experimenting with the application of innovative statistical methodologies to improve the efficiency and/or the effectiveness of data processing. Considering the need to complement their statistical skills with the necessary additional cutting-edge AI/ML knowledge, we activated a partnership with the academia to provide guidance in the development of the first applied research projects. This "on-the-job training" has proven essential for an effective kick-off of the statistical experimentations.

At the same time, a wide training program was launched to update the statistical expertise (on this see also par. 5.1) given that new techniques and methodologies are being developed at a fast rate. In fact, we

[15] Banca d'Italia Strategic Plan for 2017-2019. "Promoting the use of innovative statistical methodologies is also part of the 2023-25 Strategic Plan.

[16] Namely, "A supervised record linkage approach for anomaly detection in insurance assets granular data", "Learning from revisions: a tool for detecting potential errors in banks' balance sheet statistical reporting", "The market notices published by the Italian Stock Exchange: a machine learning approach for the selection of the relevant one", "Institutional sector classifier, a machine learning approach" and "Imputation techniques for the nationality of foreign shareholders in Italian firms".

[17] "Application of classification algorithms for the assessment of confirmation to quality remarks".

attach high priority to building in the existing staff a critical mass of advanced statistical competence and familiarity with handling the new methodologies both via targeted training programs and by offering the opportunity to experiment with applied statistical research without constraints and in mixed professional groups that also include statistical and IT experts.

With regard of the actual topics of research in the field of statistical processing, already at the end of the first innovation cycle (of 2017-19) it became clear that AI/ML solutions could prove very useful in the fields of "data validation" and "data enrichment". Over time, a third field of research has risen to the attention of our statisticians, which is the potential for improving the "efficiency of data management processes". The potential associated with the use of such innovative methods in our data management tasks as well as the pace of methodological innovations (notably the advent of generative AI) are such that there is still a lot more that can be researched with regard to different and new business cases.

3. DATA VALIDATION

Traditional quality checks for data reported by individual RAs are based on predefined, deterministic formulae (mainly derived from the reporting regulation itself) and/or thresholds (e.g. associated to the expected evolution of phenomenon that is being analysed or to benchmark values related to other RAs with similar characteristics). They are adequate to verify whether the data comply with the qualitative and quantitative relationships among phenomena that are requested by the reporting rules and their constraints. However, as granularity increases substantially, traditional checks may not be totally efficient in detecting potential outliers. Additionally, deterministic checks are also usually based on fixed thresholds, which must be periodically reviewed by a data domain expert to adapt them to the new "trends" in the data determined by the evolution of the underlying economic outlook.

To address these shortcomings, in the current context characterised by the collection of more and more granular data, the statistical area of Banca d'Italia has explored new ways to boost its traditional DQM system of regulatory reporting by introducing also more advanced, dynamic, and powerful validation methods that employ AI/ML solutions. Through their ability to learn from data ("data driven") hence keeping a high performance over time, AI/ML models intrinsically provide a dynamic approach for time and cross-sectional data, which is particularly effective in discovering patterns especially in voluminous and different reporting schemes/models.

Based on our experience, AI/ML models are more resilient to minor shocks in the underlying phenomena since they can can dynamically learn from new data and improve the signal-to-noise ratio allowing for a more accurate identification of discrepancies and anomalies at data point level. Moreover, monitoring these models and fine-tuning them over time may be more efficient than maintaining and updating an entire system of deterministic rules, especially when dealing with granular datasets. All in all, such models reduce the area of manual interventions, which instead is rather time-consuming with traditional, rigid rule-based systems.

In the data collection and processing area we addressed the problem of improving data validation in the following four works[18]:

[18] It is worth pointing out that some of the works summarised in this paper could in fact be allocated to more than one of the three areas of activity identified concerning Data validation (Chapter 3), Enrichment of registers' data (Chapter 4) and Data management efficiency and automation (Chapter 5). For instance, an AI/ML solution could be implemented for the purpose of data validation but, at the same time, it may also improve the data management processes in terms of efficiency and automation. In such cases, the allocation to a specific area of activity is based on the main goal that emerges from the work.

- *Quality checks on granular banking data: an experimental approach based on machine learning,* by F. Zambuto, M. R. Buzzi, G. Costanzo, M. di Lucido, B. La Ganga, P. Maddaloni, F. Papale and E. Svezia - Banca d'Italia Occasional papers n. 547 2020
- *Learning from revisions: a tool for detecting potential errors in banks' balance sheet statistical reporting, by F. Cusano, G. Marinelli and S. Piermattei - Quality & Quantity January 2022*
- *Stacking machine learning models for anomaly detection: comparing AnaCredit to other banking datasets,* by P. Maddaloni, D. N. Continanza, A. del Monaco, D. Figoli, M. di Lucido, F. Quarta, G. Turturiello - Banca d'Italia Occasional papers n. 689 2022

A supervised record linkage approach for anomaly detection in insurance assets granular data, by E. Svezia and V. La Serra - Banca d'Italia Occasional papers n. 821 2023

In what follows, for each paper we summarise its content, with particular reference to the business problem to be solved, the goal to be achieved, the methodological solutions found, and the main results.

3.1. Quality checks on granular banking data: an experimental approach based on machine learning

Business problem: The primary challenge addressed in the study is the effective detection and management of outliers in granular banking data (at Banca d'Italia these data are, in particular, the so called "Matrice dei conti")[19]. The study focuses on the semi-annual data on debit card issuance reported by Italian banks. The current plausibility checks (the most prominent type of checks performed on the aforementioned debit card data), that are trend-based and applied at an aggregate level, often result in a high number of "false positives", leading to unnecessary revisions and potentially reducing the quality of data checks. Additionally, the absence of homogeneity in reporting patterns across different RAs further complicates the process.

Goal: The objective of the study is to enhance the DQM system by introducing a ML-based approach to develop a robust, efficient, and automated method to detect outliers in granular banking data that is capable of handling the inherent variability and complexity of such datasets. Specifically, the study aims to develop a method that can identify potential anomalies with high precision, minimizing "false positives" and reducing the reliance on expert judgment. The proposed methodology should provide dynamic, data-driven thresholds tailored to the characteristics of individual RAs and automatically update over time as new data are collected. By doing so, the approach seeks to improve the overall efficiency and effectiveness of the DQM process.

Solution: The study proposes the use of Quantile Regression Forests (QRF) to estimate acceptance regions for outlier detection in granular data. This method is chosen for its ability to model complex statistical relationships among the data and provide robust, non-parametric estimates of conditional quantiles. The QRF model is trained using historical data on debit card issuance and includes variables that capture bank-specific characteristics and reporting patterns. The model generates prediction intervals for the target variables (number of debit cards issued), which are then used as benchmarks to identify potential outliers. These prediction intervals are dynamically updated as new data are reported, ensuring that the thresholds remain relevant over time. The empirical analysis involves dividing the dataset into training and test sets, estimating quantile functions for various levels of probability, and validating the outliers identified by cross-checking them with the RAs.

[19] Banca d'Italia Circular No. 272 of 30 July 2008 (only in Italian).

Results: The implementation of the QRF model demonstrates significant improvements in the detection of outliers compared to the traditional trend-based plausibility checks on aggregated data. The study finds that the machine learning-based approach is able to identify new anomalies that were not detected by the current DQM system, with a high level of accuracy in terms of minimizing "false positives". Specifically, the model identifies potential outliers in the number of debit cards issued with a precision rate that is significantly higher than the existing system's performance. The findings indicate that the proposed methodology can complement the current DQM system by providing an additional layer of checks on granular data. Moreover, the automated and purely statistical nature of the new approach makes it less time-consuming and more suitable for large datasets. The study concludes that machine learning techniques offer a promising solution for enhancing DQM processes in central banks.

3.2. Learning from revisions: a tool for detecting potential errors in banks' balance sheet statistical reporting

Business problem: The challenge is to detect possible errors in the Italian banks' balance sheet statistical reporting, which is the source used to calculate the national contributions to the Eurosystem's credit and monetary statistics[20]. However, the diagnostic checks based on fixed thresholds are not tailored to the reporting behaviour of each bank. As a result, potential errors detected by these methods need to be individually reassessed, leading to a time-consuming and labour-intensive process.

*Goal***:** The objective is to use AI/ML methodologies to develop a more efficient and accurate automated procedure to identify potential errors in banks' balance sheet statistical reports by exploiting historical reporting behaviour of each bank. The focus is on the outstanding amounts of loans to non-financial corporations and households. The error detection mechanism aims to enhance the quality of the Italian component of Eurosystem loans statistics.

Solution: The proposed solution is the development of a Revisions Adjusted – Quantile Regression Random Forest (RA–QRRF) algorithm. This machine learning supervised approach leverages the extensive historical data on banks' reporting errors and revisions to calibrate the acceptance regions for reported values. The RA-QRRF algorithm predicts the intervals within which the reported values should fall, adjusted by a monthly "imprecision rate" specific to each bank. This rate is calculated from each bank's entire history of reporting errors and revisions. The imprecision rate further refines the prediction intervals, penalising banks with a history of frequent errors and rewarding those with fewer mistakes. This approach allows for real-time error detection, as the necessary balance sheet variables are available during the BSI production round.

*Results***:** The empirical results demonstrate that, compared to traditional methods, the RA-QRRF algorithm significantly improves error detection as well as reduces the number of false positives. Despite the higher variance and more complex reporting behaviour associated with corporate loans, the RA-QRRF outperforms existing methods and provides a substantial improvement in error detection accuracy. This machine learning-based approach offers a robust, scalable solution for enhancing the quality and reliability of banks' balance sheet data, supporting the ECB's monetary policy and financial stability objectives. In a nutshell, the RA-QRRF algorithm represents a significant advancement in data quality management practices at Banca d'Italia, leveraging machine learning to efficiently and accurately detect potential errors in statistical reports. This innovation not only improves the accuracy of monetary statistics

[20] National Central Banks regularly collect (and disseminate) monthly data on monetary and financial institutions balance sheet items (BSI) according to the *ECB Regulation ECB/2013/33 of 24 September 2013 on the balance sheet of the monetary financial institutions sector* and the *ECB Guideline of 4 April 2014 on monetary and financial statistics (ECB/2014/15)*. https://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX%3A32013R1071
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32014O0015
Please note that in 2021 the ECB replaced Regulation ECB/2012/33 with Regulation ECB/2021/2.

but also streamlines the data validation process, reducing the reliance on manual intervention and enhancing the overall efficiency of the BSI production process.

3.3. Stacking machine learning models for anomaly detection: comparing AnaCredit to other banking datasets

Business problem: Ensuring the quality of the AnaCredit[21] survey on "individual" loans to legal entities poses significant challenges. Traditional deterministic methods to cross-check AnaCredit aggregates with similar aggregated information as those under the "Balance sheet items"[22] (BSI) and "FinRep"[23] frameworks show limitations as the rules rely on fixed acceptance thresholds that do not account for variability across banks and time periods. This approach fails to identify context-specific anomalies and can result in either an excessive number of "false positives" or overlooked errors. The problem is compounded by the structural differences between the three datasets, mainly with regard to the reporting thresholds and the definitions of credit instruments. A manual verification, which could help in introducing some flexibility in the verification, is not only time-consuming but may also lead to inconsistent decisions since it heavily relies on the subjective judgment of data managers. This uneven treatment can lead to uneven data quality across different reporting periods and reporters, undermining the overall reliability of the data.

Goal: The objective of the study is to develop a ML-based framework that can automatically detect anomalies in the aggregates built with individual AnaCredit data by comparing it with the loan aggregates derived from BSI and FinRep datasets. This framework aims to leverage the renowned high-quality data from BSI and FinRep. By employing advanced statistical and machine learning techniques, the model aims to establish a systematic cross-checking mechanism that can handle the complexity and granularity of the AnaCredit data. The ultimate goal is to make the data quality framework more effective by enhancing it with robust cross-checks with alternative reliable sources while reducing the possible manual double-checks burden on both the data managers and the RAs.

Solution: The proposed solution involves an "ensemble method" combining Robust Regression and Autoencoder models within a semi-supervised learning framework. The study employs a robust regression model to account for the structural differences between AnaCredit and the benchmark datasets, BSI and FinRep. This model aims to identify and isolate reporting errors by capturing the relationship between the datasets while considering their inherent differences. Simultaneously, two Autoencoder models - a Convolutional Autoencoder (AE-CNN) and a Dense Autoencoder (AE-DNN) - are also used to detect anomalies. These models reconstruct the input data, after embedding it in a lower-dimensional space, and measure reconstruction accuracy using the Structural Similarity Index Metric (SSIM), which is designed to capture perceptual similarity. The final step involves combining the outputs of these base models through a stacking algorithm. This meta-classifier is trained on the binary labels "anomalous" or "not-anomalous" derived from domain knowledge and cross-checks with RAs on only a subset of the overall available data. The use of semi-supervised learning allows the meta-classifier to leverage both labeled and unlabeled data, enhancing its predictive performance. The overall process ensures that the anomalies identified are contextually relevant and can be effectively addressed by the RAs.

[21] *ECB Regulation (EU) 2016/867 of 18 May 2016 on the collection of granular credit and credit risk data (ECB/2016/13)*, so-called AnaCredit Regulation. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016R0867

[22] Please refer to footnote 17.

[23] *Regulation (EU) 2021/451 17 December 2020 laying down implementing technical standards for the application of Regulation (EU) No 575/2013 of the European Parliament and of the Council with regard to supervisory reporting of institutions, and repealing Implementing Regulation (EU) 680/2014*. https://eur-lex.europa.eu/legal-content/it/TXT/?uri=CELEX%3A32021R0451

Results: The empirical evaluation of the proposed models demonstrates a significant improvement in anomaly detection compared to traditional methods. The Robust Regression model effectively captures the structural differences between AnaCredit and the benchmark datasets, while the Autoencoder models prove themselves to be very effective in detecting subtle anomalies. The meta-classifiers outperform individual base models on the fully-labelled dataset, providing a more reliable final prediction. The findings show that the stacking technique improves the F1-score, a metric suited for evaluating the performance of models in the presence of imbalanced data. To get the fully-labelled dataset, two different approaches are considered and for each of them different meta-classifiers are trained. In the first approach, true labels are assigned to the entire set of observations using a Monte Carlo simulation: according to the benchmark dataset, the Support Vector Machine (SVM) model or the Random Forest model perform best. In the second approach, semi-supervised models are trained using only the sampled and pre-labeled variable, with the labeling occurring during the model estimation process. In getting the fully labelled dataset, either the Co-Bagging algorithm or the self-training algorithm yields the best results according to the benchmark dataset. Overall, the semi-supervised approach outperforms the Monte Carlo simulation. These findings suggest that the proposed methodology can significantly enhance the quality and the accuracy of control processes for AnaCredit data. The framework developed in this study is flexible and can be adapted to other datasets with similar characteristics, providing a generalizable approach to improving data quality through machine learning techniques and allowing tailoring the models to the needs at hand.

3.4. A supervised record linkage approach for anomaly detection in insurance assets granular data

Business Problem: This study addresses the issue of erroneous changes in asset Identification Code (ID code) reported by insurance corporations. Since 2016, insurance companies have been required to submit detailed asset data in Solvency II templates on a quarterly basis. Each asset is identified by a unique ID code that should remain stable and consistent over time. However, due to reporting errors, these codes can change unexpectedly, leading to inconsistencies in insurance statistics. Such inaccuracies might suggest that an asset has been removed or added to an insurer's portfolio, resulting in misleading interpretations of the time series in insurance statistics.

Goal: The objective of this study is to use AI/ML to develop a methodology to detect anomalies in reporting the ID code of insurance assets. By employing a supervised machine learning approach within a statistical matching framework, the study aims to accurately and efficiently identifying changes in asset ID codes that can be indicative of reporting errors. The ultimate goal is to improve the overall quality of the data, thus enhancing the reliability of the statistics compiled by central banks and supervisory authorities. This approach also seeks to minimize the burden of manual verification on data analysts and streamline the DQM process.

Solution: To address the above-mentioned problem, the study proposes a machine learning-based record linkage approach to detect anomalies in the reporting of insurance asset ID codes. The methodology involves several steps, starting with the construction of a comparison matrix where each pair of assets from two adjacent quarters is compared based on a set of qualitative and quantitative features. These features include nominal variables (e.g., counterparty sector), ordinal variables (e.g., categorized maturity date), numerical variables (e.g., market value), and textual variables (e.g., asset description). Comparisons are conducted through the computation of appropriate distance measures, such as the overlap measure for nominal variables and the Levenshtein distance for textual variables. The resulting comparison matrix serves as input to supervised classification models, with the binary target variable indicating whether each pair of assets is a match or a non-match, i.e. if their ID code is equal or not.

The matrix is then split into a training and a test set, stratified by asset type and sampled to be unbalanced with respect to the target variable, for different imbalance proportions.

Three classes of models are evaluated: Logistic Regression, Random Forest and Neural Network. The Random Forest model, known for its robustness and high performance in various domains, is selected based on its superior performance in terms of accuracy and balanced accuracy across different test scenarios.

Results: The empirical findings demonstrate that the proposed Random Forest model can efficiently detect changes in the reported asset ID codes, with high performance results, robust to changes in the tested date (i.e. different pairs of reporting quarters) and to differently unbalanced datasets. Indeed, the model ensures high "true positive" rates with "false discovery" rates across various test scenarios that are acceptable for keeping the DQM process efficient; also the average accuracy and balanced accuracy are set at very high levels. Additionally, specific thresholds for classification probabilities are fine-tuned for different asset types to further enhance performance.

In real production rounds conducted in 2022, analyses have confirmed that around 60% of the cases identified were indeed anomalies. These results highlight the model's potential to streamline DQM processes and reduce manual verification efforts significantly. The study concludes that the machine learning-based approach not only enhances the reliability of insurance statistics but also opens new avenues for future research and improvements in DQM methodologies.

4. ENRICHMENT OF REGISTERS' DATA

Data estimation is a viable strategy not only for data analysis and forecasting but also in situations in which central banks want to close informative gaps without imposing an excessive burden on RAs. Registers' data are a key ingredient in the data collection process to build sound aggregate statistics from very granular data, so it is important to replace possible missing values with robust estimates. At Banca d'Italia, the statistical function takes care of this task by centralising its management so that the same unique estimations are used in all calculations and also users are lifted from such pre-processing activities.

Against this background, AI/ML offers a promising approach to efficiently deal with missing data imputation, since such techniques allow to rapidly analyse and process large volumes of data and identify the patterns and relationships among them. In turn, such techniques lead to more accurate predictions for the missing values than with traditional statistical models such as "deletion" or "mean imputation"[24].

In the data collection and processing area the imputation of missing values without imposing an additional burden on RAs was tackled in two works:

- *Institutional sector classifier, a machine learning approach* by O. Giudice, P. Massaro and I. Vannini - Banca d'Italia Occasional papers n. 548 2020
- *Imputation techniques for the nationality of foreign shareholders in Italian firms* by A. Carboni and A. Moro - IFC Bulletin n. 48 2018

Below, for each paper we summarise its content, with particular reference to the business problem to be solved, the goal to be achieved, the methodological solutions found, and the main results.

[24] "Deletion" reduces the sample size, potentially weakening the statistical power of the analyses; "mean imputation" requires assumptions that are not always valid given the underlying process that causes some data to be missing

4.1. Institutional sector classifier, a machine learning approach

Business problem: With the increased collection of "individual data" related to businesses and firms, the attribute "institutional sector"[25] stored in the Bank's Entities Register is used to build meaningful economic statistics as it allows the aggregation of data pertaining to the individuals classified in the same sector. In this regard, Banca d'Italia defines its own classification (called Sector of Economic Activity, SEA[26]) which draws heavily from the European System of National and Regional Accounts (ESA 2010[27]), though it offers more breakdowns.

In the Entities Register different sources are used to assign the SEA code to each entity depending on its type, specifically:

- the same Banca d'Italia, for supervised institutions;
- the national statistical institute (ISTAT), for the public sector;
- the insurance supervisory authority (IVASS), for insurance companies;
- RAs, for the remaining entities.

Evidence shows that RAs are not always able to report the SEA code of an entity. Or, in other cases, SEA codes provided by RAs have a certain degree of inaccuracy. Among all RAs, this problem is mostly significant for financial corporations.

Goal: The primary objective is to develop a predictive model of entities SEA code merging the information on the entities that can be derived from different available sources (e.g. the very denomination of the entity as it appears in the Bank's Entity Register, Italy's Revenue Agency data base, the National Business Register, the Ministry of Economy and Finance data base). This solution aims to achieve two key goals:

- Imputation: Accurately estimate missing values for the "Institutional sector" attribute, ensuring a complete dataset for further analysis.
- Validation: Enhance data quality by automatically verifying updates submitted by reporting agents. This would free up human resources currently dedicated to manual data validation.

Solution: The study first focuses on identifying a set of predictors that effectively classify entities within their respective sectors. This involves pre-processing both structured and unstructured data, selecting candidate machine learning algorithms, and employing a validation step to choose the best performing model. A significant challenge concerns addressing the unbalanced nature of the data. Most of the entities typically belong to the non-financial sector, with a smaller portion distributed across various financial sectors. This imbalance poses the risk of the null model, assigning all entities to the non-financial sector, potentially outperforming a machine learning model.

The proposed solution involves applying a series of Gradient Boosting models sequentially in a hierarchical order to ensure balanced class distribution at each step. However, this approach relies on a pre-defined hierarchy based on business knowledge, raising concerns about its limitations.

To address these concerns, the hierarchical approach is compared with a neural network model that did not consider a predetermined hierarchy. Additionally, various scenarios combining these approaches are evaluated against the null model, using error rate as the performance metric.

[25] The institutional sector characterises a company's primary function and its economic role.

[26] Banca d'Italia classification of the Sector of Economic Activity (SEA) is described in Circular No. 140, 11 February 1991.https://www.bancaditalia.it/statistiche/raccolta-dati/segnalazioni/normativa-segnalazioni/c140/circolare_140_ottobre_2021.pdf, available only in Italian.

[27] The ESA 2010 is the European coding system developed by EUROSTAT to classify economic units.

Results: The implementation of the models in various scenarios, including the hierarchical and neural network approaches, demonstrates similar levels of accuracy. Therefore, to optimize efficiency, the simplest model with a satisfactory level of accuracy is chosen for integration. This streamlined approach effectively reduces the burden on RAs while maintaining a high degree of accuracy. The solution has been successfully integrated in 2021 into the regular data collection process, reducing the time required for data validation and ensuring a complete and reliable dataset for further analysis.

4.2. Imputation techniques for the nationality of foreign shareholders in Italian firms

Business problem: To estimate the Foreign Direct Investments (FDI), Banca d'Italia relies on the information provided by a sample of non-financial and insurance companies, where the inward and outward FDI relationships of the firms are also collected. Information on "FDI inward" is generally available in the Italian Chambers of Commerce, where companies report the list of their shareholders, however, most of the time the nationality of the shareholders is missing.

Goal: The purpose of this study is to exploit ML methodologies to identify the firms' nationality when unknown, and the only relevant available information is the denomination of the corporation.

Solution: To solve this problem, the authors develop an imputation algorithm for the country of residence of an entity based on its denomination. Firstly, dummy variables are created for each word in the denomination (bag-of-words approach); then, a statistical model is estimated linking the nationality of each firm to the dummy variables. In this regard, as the data is very unbalanced – the vast majority of the shareholders are Italian companies – a first Logistic Regression model distinguishes Italian and non-Italian companies, and for the latter, a second Multiclass Logistic Regression model is estimated to obtain the country of residence.

Results: The out-of-sample analysis shows that the accuracy of the algorithm is very high (around 98%), with an almost perfect discrimination between Italian and foreign firms. Moreover, the proposed approach is able to classify correctly most of the countries with high levels of foreign direct investment in Italy.

The proposed model was evaluated against a Random Forest, concluding that the classification performance of the two-step algorithm is competitive with respect to other non-parametric ML methods.

5. DATA MANAGEMENT EFFICIENCY AND AUTOMATION

The automation of some areas of data management (through statistical and IT solutions) is motivated by the need to reduce manual and repetitive interventions so as to devote staff to new activities. In particular, the focus is on processing large amounts of data at a much faster rate than a human can, reduce human errors and increase objectivity and standardization when handling different data sets. In this scenario, human contribution to the overall process has remained key for what concerns data analysis (e.g. evaluation of outliers, comparison of results over time and across RAs), communication with RAs and decisions on the overall fitness-for-use of a dataset in the presence of pending anomalies. The advent of AI/ML solutions opens new and unprecedented possibilities to further improve the efficiency and effectiveness of the data management processes for their capabilities of emulating human decision skills and incorporating past experience so as to maintain (or even improve) the appropriateness of the decisions over time. Human monitoring obviously remains key, but delegating more and more steps of the workflow to an automatic procedure makes it possible to devote highly skilled staff to new relevant activities.

In the data collection and processing area we analysed the issue of increasing, through ML methods, the efficiency of the reporting management process in three papers:

- *Application of classification algorithms for the assessment of confirmation to quality remarks*, by F. Zambuto, S. Arcuti, R. Sabatini and D. Zambuto - Banca d'Italia Occasional Papers n. 631 2021
- *A decision-making rule to detect insufficient data quality: an application of statistical learning techniques to the non-performing loans banking data,* by B. La Ganga, P. Cimbali, M. De Leonardis, A. Fiume, L. Meoli and M. Orlandi - Banca d'Italia Occasional Papers n. 666 2022
- *The market notices published by the Italian Stock Exchange: a machine learning approach for the selection of the relevant ones*, by M. Bernardini, P. Massaro, F. Pepe and F. Tocco - Banca d'Italia Occasional Papers n. 632 2021

  Again, below for each paper, we summarise the contents of the works with reference to the business problem to be solved, the goal to be achieved, the methodological solutions found, and the main results.

5.1. Application of classification algorithms for the assessment of confirmation to quality remarks

Business problem: In the context of the DQM of supervisory banking data, analysing the motivations received from RAs on quality remarks in order to decide if a confirmation can be accepted or refused is a critical step in ensuring a high standard of quality and improving the overall DQM system. Such a stage of the DQM process is a manual task assigned to the Data Manager that can be highly time-consuming and subject to various inefficiencies, specifically in the presence of more complex surveys, since:

- the number of potential outliers can be quite large;
- the cases of "false positives" can become recurrent in the system when the validation checks rely on assumptions that are not valid for all plausible reporting patterns; in turn, a reporting exception, although already known, can affect the data submitted by several RAs;
- the process requires some degree of judgment by Data Managers (potentially resulting in heterogeneous patterns if similar cases are treated differently over time and across data collections and RAs).

Goal: The purpose of the paper is to verify if AI/ML techniques can reduce time and limit human discretion and misjudgment in processing the confirmations to the data provided by the RAs. In particular, the authors explore the use of AI/ML techniques to automatically recognize known cases of admissibility or ineligibility of confirmations, with the ultimate goal of establishing a more efficient process while ensuring a higher data quality.

Solution: An innovative methodology is proposed for the automatic processing of the data confirmations. The solution is based on text mining and a supervised learning algorithm. Specifically, a classification model is trained to replicate the decision-making process of data managers with regard to the acceptance (in the case of "false positives") or rejection (in the case of true positives) of confirmations. The decision rules take into account both the textual explanation provided by RAs and the characteristics of validation checks, as well as the overall reporting behaviour of the whole banking system.
A Random Forest and a Gradient Boosting model were implemented and, between the two, the Gradient Boosting was selected. Regarding the model selection, it is worth pointing out that the authors put a lot of attention to the 'user experience' of their solution. Moreover, when choosing the solution, it proved crucial to take into consideration the possible consequences of producing too many "false positives", in terms of the risk of determining distrust in RAs and of the cost of the anomaly analysis process. Lastly, the interaction with data managers was needed to interpret the model results (even in the absence of errors in the data) and to explain the detected anomalies.

Results: The empirical findings show that the methodology predicts the correct decisions on recurrent data confirmations and that the performance of the proposed model is comparable to that of Data Managers, outperforming them in some cases.

5.2. A decision-making rule to detect insufficient data quality: an application of statistical learning techniques to the non-performing loans banking data

Business problem: Usually data quality follows a positive trend thanks to subsequent corrections submitted by RAs; however, a worsening of quality may indeed occur, especially when data production is affected by exogenous and unpredictable events such as issues in the RAs' IT tools, recent changes in the reporting requirements that RAs have not completely mastered or process failures (also due to unforeseen staff shortages, as during the pandemic).
The problem is to establish, for a given RA and reference date, the quality variation between two subsequent reporting submissions when they both have some quality issues but do not bear "serious" remarks (which would otherwise block the dissemination) and therefore are relatively fit for use. In the current practice, the comparison between two consecutive datasets submitted by the RA is left to the judgment of the data manager who decides upon their expertise and the explanations possibly received by the RA.

Goal: The purpose of the study is to define a decision-making rule, based on AI/ML techniques, to automatically evaluate if revisions improve or worsen the data quality level (DQL) of a dataset when its quality is affected only by "non-serious" issues. This is a sort of "grey area" that is usually manually assessed by the data manager who also must face a trade-off between the need to make the information promptly available to users, and the need for the data to be "fit for use'" i.e. free from significant errors. An automated algorithm could save time and allow for a more agnostic approach to data quality by not considering the individual sensitivities.

Solution: The study evaluates the application of statistical learning techniques that, starting from the actual evidence of previously reported datasets, can provide, as a first step, a prediction on whether a remark would be confirmed or would trigger a revision by the RA. Specifically, a Logistic Regression model is selected to predict the confirmation probability. Such a prediction is subsequently used to estimate the probability that confirmable remarks are indeed confirmed. This allows for the identification of two groups of observations in a dataset: the first is composed of those reports whose DQL is greater than the previous one sent by the RA, the second contains those reports whose DQL is lower and are then likely to be revised. The implementation of this statistical technique leads to a synthetic data quality indicator that, together with the decision-making rule, provides guidance on the overall fitness for use of the information contained in the dataset.

Results: The findings show that a decision-making algorithm could appropriately support the data manager in deciding whether the new data have a sufficient DQL and can then be disseminated. To verify the appropriateness of the decision rule to classify further submissions as "to be disseminated" and "to be blocked", a comparison with a benchmark is carried out. In particular, the benchmark is obtained by applying the decision rule when the observed variable "The remark is confirmed Yes/No" is considered instead of estimation according to the Logistic model. The empirical findings show that in 97% of cases, the decision would be the same as when the actual status was known.

The rule may help the data manager to determine whether data revisions do improve the data quality of previously reported data and to distinguish the reports that most likely will need to be corrected from those that will instead be confirmed by the RA. The decision-making rule also provides guidance to data managers for prioritizing data quality activities for the identification of insufficient data quality reports

that require a new submission from the RAs. At the same time, it provides a tool to identify reports that may not be immediately suitable for use.

5.3. The market notices published by the Italian Stock Exchange: a machine learning approach for the selection of the relevant ones

Business problem: the Italian Stock Exchange - *Borsa Italiana* - regulates the procedures for listing companies and supervises the disclosure of related information through the publication of Market Notices every year on its website, around 25.000 in 2020 of which around 4.000 include information useful for the Bank's Securities Database[28]. Therefore, on a daily basis, the data manager has to check an average of over 100 published notices; it is an activity that not only is highly time-consuming, but also error prone to the extent that relevant notice might be overlooked.

Goal: The purpose of this study is to use ML tools to automate the selection of the Market Notices containing relevant information to update the Securities Database, for example, the results of an offer of change in capital.

Solution: A supervised model is trained to perform the binary classification task "relevant"/"non-relevant". Before the ML model is applied, the information from the first page of the Market Notices, which contains semi-structured data, is prepared and vectorised with the "bag-of-words" technique. Then, the binary classification model is trained and fine-tuned to predict the categorical variable Y, which takes two values (*Yes=1* and *No=0)* depending on whether the notice is of interest or not. Different models are tested, such as the Naïve Bayes, a Random Forest, and the Logistic Regression. Among the three fine-tuned models, the Logistic Regression achieves the highest weighted accuracy score. Defining a robust measure and evaluating the overall quality of a model with a single performance measure speeds up the model selection and the hyperparameter fine-tuning process. With regard to the model selection, it is crucial to identify the most relevant metric(s) to the business goal. Moreover, the detection of the most appropriate cost function to be optimised required very close business-IT interactions.

Results: Empirical findings show that, despite its simplicity, the final Logistic Regression model outperforms the manual approach of the data managers. Such a procedure also presents the advantage of ensuring a uniform classification over time and across different data managers, hence removing the risk of a potential bias due to the subjective selection made by each individual data manager. The importance of the analysis is twofold: firstly, it increases the overall quality of the database; secondly, it enhances the efficiency of the whole statistical production process in that a highly time- and resource-consuming activity was carried out automatically. It is also worthwhile remarking that the proposed methodology is strictly dependent on the format of the Market Notices published by the Italian Stock Exchange (on this see also chapter 6).

6. LESSONS LEARNT

The importance of the results achieved in the studies described above go beyond mere methodological findings, offering valuable insights into the factors that enable innovation in the data production process. Over time we have implemented in the regular operational processes five of the studies summarised in this paper, and a sixth is about to be released[29]. Focusing on the remaining three works that have not (yet) been implemented, we observe that the main reason for them not to have been realised traces back to the difficulty to reallocate staff from current activities to the innovative ones; as such, it can be regarded as a temporary factor, which could be overcome in the medium term through a reprioritisation of activities.

[28] Please refer to footnote n. 12.

[29] Please refer to footnotes 16 and 17 for the details.

Below the main lessons learnt so far are summarised with the aim of highlighting the conditions that foster the successful implementation of AI/ML solutions in statistical processes related to regulatory reporting. Such lessons are presented alongside potential *ex-ante* and *ex-post* measures to address the challenges that can be encountered throughout the process.

6.1. The human capital

In moving from the study of an AI/ML solution to its realisation, deployment, running and maintenance, the very first challenge is related to the availability of human resources with cutting-edge statistical skills that are essential to adequately deal with such innovative methodologies. This is particularly critical in the field of AI/ML, where training on an ongoing basis is crucial, due to the rapid evolution of methodologies which requires continuously updating the knowledge acquired at the university or in previous learning activities or work experiences.

To address this issue, aside from the usual internal courses on data management and statistics, the statistical function at Banca d'Italia has sponsored a threefold training package in data analytics, then implemented in cooperation with academia and the HR Directorate, which comprises (i) a full II level MS degree in Data Science, (ii) targeted post-graduate courses selected therefrom, (iii) an internal Data Science Academy[30]. These tuition paths - that are usually alternative to each other - are offered to selected, highly motivated and talented professionals already in the ranks of the statistical function. In our experience, an advanced statistical training proved crucial for supporting our applied research and helped us to keep pace with the constant evolution of AI/ML technologies and methodologies together with the parallel hiring of Data Scientists[31]. Besides, the very first group of research projects presented in this paper, also benefited from dedicated courses, provided by university professors, which were tailored to the methodological knowledge required for each project ("training on the job"). Such support proved fundamental to rapidly acquire the necessary skills to transform ideas into methodologically sound papers and prototypes.

Collaboration with academia, however, may not suffice given the multitude of possible areas of application to improve the efficiency and effectiveness of the regulatory reporting management process and the costs and complexity of such an investment. Considering that, in this field, the business cases may be similar across central banks, a collaboration could be established to pull together human and IT resources so as to tackle common business cases and, therefore, benefit from economies of scale and synergies[32]. An important initiative along this direction is the "Virtual Lab" provided by the European Central Bank, i.e. a web-based platform providing collaboration and innovation capabilities, including the use of generative AI[33]. Also, the Irving Fisher Committee on Central Bank Statistics[34] (IFC) is very active in organising events and workshops where central banks can exchange views and experiences as well as create

[30] The Data Science Academy is a two-year training program - specifically tailored for Banca d'Italia - comprising both theoretical classes and hands-on activities closely related to actual business cases with the aim to shape a distinct professional profile.

[31] Fostering an in-house development of the necessary skills has proven successful to compensate, on the one hand, the difficulties of a timely acquisition of the right talent considering that Banca d'Italia must hire its staff by means of public competitions and, on the other, the need to compete with a vast and diversified private sector to attract the most highly qualified professionals (i.e. PhD students). Besides, the constant evolution of AI/ML techniques requires a continuous learning process that would be necessary even when hiring dedicated profiles.

[32] When feasible, also partnerships with the private sector - especially start-ups - could prove particularly fruitful as such companies are often at the forefront of technological innovation.

[33] On the ECB Virtual Lab, please refer to E. McCaul speeches at the Supervision Innovators Conference in 2022 and 2024.

[34] The IFC is a forum of central banks economists and statisticians established and governed by the central banking community operating under the auspices of the Bank of International Settlements.

opportunities for shared work[35]. However, in a highly diversified and complex context, where authorities operate in different jurisdictions, so far, the role held by the above initiatives has proven insufficient to promote the concrete realisation of common statistical innovations and transform what is essentially a network for information exchange into a comprehensive and mature functional framework for developing AI/ML solution for data collection and processing. A major step forward would be related to a clear commitment from the authorities involved to pooling human resources on a common work programme; this initiative should go be accompanied by the development of a common IT platform suitable for carrying out joint AI/ML projects.

Given the sophistication of AI/ML models, two other relevant aspects must be taken into account in the management of high-skilled staff: how to deal with the staff turnover and how to bring together the required diversified expertise.

Turnover may pose a significant challenge where there is a limited degree of substitutability of the skills. To ensure that junior data managers can quickly become proficient with existing models and processes, it is crucial to plan an effective take-over and knowledge-transfer of the model and the process, while also fostering a standardisation of approaches (as and when possible) and rewarding teamwork so that the relevant know-how is not confined to a few. These measures are necessary for a thorough understanding of the inner workings of the models to appropriately interpret the results and take care of their ongoing maintenance. It is also worth mentioning that a thorough documentation - of the model, the deployed solution and the related operative procedures - proves vital in presence of turnover within the team that developed the AI/ML solution, in order to effectively manage breaks in the data or a deterioration in the performance of the model.

Regarding the expertise of the people involved in designing and developing AI/ML solutions, pooling diversified professional skills is a success factor for the development of effective and sound solutions. Indeed, establishing multidisciplinary teams triggers fruitful discussions and fosters brainstorming of ideas, and this is why it is important to bring together people with different background such as statisticians, data scientist, computer scientist as well as people expert of the business case[36]. In such teams business experts provide insights for a correct interpretation of the results and, from a methodological point of view, a pivotal role is played by the data scientist whose professional competence covers several disciplines and therefore is able to connect the different perspectives. For the same reasons a multidisciplinary approach is proven instrumental also for the monitoring and maintenance of an AI/ML application, among other things since it facilitates knowledge sharing and allows for an easier handling of temporary skill gaps.

To sum up, in this new rapidly evolving statistical and technological landscape, it is important to have the right skills during the whole lifecycle of an AI/ML solution. To this end, recruiting the people with the right skills may not be sufficient, since the statistical knowledge rapidly becomes obsolete. Therefore, training plans to keep up with the methodological evolution and fostering the micro-knowledge transfer among the developing teams' members are all key success factors.

6.2. Model identification and management

[35] Machine learning applications in central banking", by D. Araujo, G. Bruno, J. Marcucci, R. Schmidt, B. Tissot, IFC (2021) and "Data science in central banking: applications and tools", by D. Araujo, G. Bruno, J. Marcucci, R. Schmidt, B. Tissot, IFC (2022).

[36] Depending on the complexity of the AI/ML project, some roles could be in fact played by the same individual; this is, for example, the case of most of the studies illustrated in this note where the same person plays the role of the statistician as well as of the business expert.

Once the multidisciplinary teams with the right skills have been formed, further issues that must be addressed relate to the identification and management of the model.
As a first consideration, it is essential that the complexity of the model does not exceed the one of the process that is generating the data so that the model can appropriately capture the data pattern(s), preventing overfitting and ensuring reliable predictions. For example, considering AI/ML applied to data quality checks, an overly complex and sensitive model that captures noise instead of meaningful patterns may generate too many "false positives". This has practical consequences since it may cause desensitisation in banks due to an excessive number of warnings.
Moreover, once an AI/ML solution is implemented, it usually requires a periodic recalibration or even a redesign to consider possible drifts in the original data[37]. The evolution of the model may be simply triggered by changes in the format of the source data. These activities are an integral part of the adoption of AI/ML solutions if you want to ensure that the model and the process remain aligned with evolving data landscapes and maintain their performance level and practical relevance.
For example, after two years since the deployment of the application to automatically select the Market Notices (see par. 4.3), *Borsa Italiana* changed the format of the Market Notices causing a large drift in the model[38].

6.2.1. Ethical implications

With reference to the selection of the most suitable model and its specification, another important aspect regards the "responsible use of AI/ML", an ethical category that, for the scope of this paper, unfolds into (1) the issues of energy consumption and (2) the difficulty to ensure transparency of the results.

Regarding the first aspect, attention should be paid to the amount of electricity consumed by an AI/ML model, that has both economic and environmental implications. This is a novelty that requires a paradigm and cultural shift on the part of AI/ML model developers, who must become conscious of the energy footprint of their models and the consequent importance to take it into account in the model's development process. Furthermore, also the maintenance of AI/ML models is energy-intensive, and this must be factored into the overall energy consumption equation [39].

The second aspect highlighted above is particularly relevant for central banks, although it can be inherently quite challenging in the context of AI/ML. In fact, just as it is necessary to ensure **transparency** and be accountable in traditional statistical production, it is equally clear within the official statistics community (which also includes central banks) that these ethical categories must also be adopted when AI/ML methods are used[40]. Instead, AI/ML models can be complex and difficult to interpret, leading to challenges in explaining how statistical conclusions are reached, especially when employing the latest generative models[41]. Ensuring transparency is not always an easy task since the model quite often appears

[37] We refer, for example, to changes in the content of external data sources due to a new regulatory framework, the adoption of a different technical model to design the same data source (even just a reshuffle of the format of the data file) or classification modifications.
[38] At the same time more structured information became available on *Borsa Italiana*'s web site, making it possible to select the relevant notices by using deterministic rules. Thus, in this circumstance, it was not necessary to update the model.
[39] "Precise energy consumption measurements of heterogeneous artificial intelligence workloads" (2022), by R. Caspart, S. Ziegler, A. Weyrauch, H. Obermaier, S. Raffeiner, L.P. Shumacher, J. Scholtyssek, D. Trofimova, M. Nolden, I. Reinartz, F. Isensee, M. Goetz, C. Debus. ISC Workshops 3 Dec 2022.
[40] UNECE, "Machine Learning for Official Statistics", 2022
[41] UNECE, "Large Language Models for Official Statistics", 2023

as a "black-box"[42]. In the field of reporting, this is particularly relevant for the outcomes of data quality checks, since RAs should be given the opportunity (and the information) to easily find the underlying cause for the potential outliers detected by an AI/ML model, also to prevent its repetition in the future. Moreover, for regulatory reporting where data anomalies may trigger non-compliance interventions of the relevant authority, it is of paramount importance for the latter to be able to transparently demonstrate the irregularity in the data, so as to counterbalance its power to request corrections. It is worth remarking that there is often an inherent trade-off between transparency and the ability of the AI/ML model to identify potential reporting errors. Advanced AI models capable of processing and analysing complex data tend to rely on deep sophisticated mechanisms to uncover hidden learning patterns in place of explicit symbolic expressions or predefined logical rules[43]. While this enables them to achieve superior performance and detect subtle anomalies, it significantly challenges their explainability capacity, making it harder to trace or justify their decisions. This trade-off underscores the delicate balance between maximizing model efficacy and maintaining the transparency needed for accountability and trust.

When the data management process has a direct impact on RAs, as in the case of DQM, the statistical function of Banca d'Italia has opted for a strong "human-in-the-loop" approach. In fact, the results of the quality checks produced by an AI/ML solution are not forwarded automatically to the RAs, rather they are analysed by the data manager and then communicated with the side information to allow them to better understand the anomalies found in the data (also committing to the possibility of further explanatory exchanges). Avoiding fully automated processes is somewhat time-consuming, but it is a necessary measure to ensure that the results are accurately evaluated and appropriate feedback to the industry is provided. For example, there are cases in which AI/ML models benefit from a preliminary segmentation[44] of the units (e.g. when the reporting behaviour is compared with that of "peer" RAs) and others in which said segmentation is intrinsically operated by the very same model (e.g when the reporting bahaviour of a RA is correlated with its past behaviour as in the paper "*Learning from revisions: a tool for detecting potential errors in banks' balance sheet statistical reporting*" discussed in par. 3.2). Either way, it is necessary to ensure that a segmentation is justified by the quality of the results and does not produce unreasonable bias (in terms of unjustified pressure to solve data quality issues) towards the RAs themselves. That is when the scrutiny of the "human" is crucial to ensure a fair treatment.

Another action that we consider useful to support transparency and accountability is publishing technical papers that document the methods, and the algorithms adopted (or intended to be) in the data management processes. In fact, the lack of traceability may undermine the confidence in the disseminated data and such an erosion of trust can carry significant societal implications, particularly when this information supports critical policy decisions. In this respect, all the AI/ML solutions reviewed in this paper have been presented as individual contributions published in the Bank's Occasional paper series. This has allowed to engage a scientific discussion on the matter with other organisations and the academia[45], ultimately fostering trust in the institution and shaping the scientific debate within the community of official statistics.

[42] Please consider that, as far as Banca d'Italia's regulatory reporting is concerned, applications of AI/ML solutions do not have an impact on natural persons' health, safety or fundamental rights. For this reason, the AI Act provisions (Regulation EU 2024/1689), for example about "explainability", do not directly apply. A continuous monitoring is anyway in place to ensure that, should this change in the future, compliance with AI Act requirements is respected.

[43] So called "sub-symbolic AI" of which deep learning systems used for language processing are an example.

[44] Segmentation is the process by which statistical units are divided up based on common characteristics. In the case of regulatory reporting, RAs could be grouped by localisation, type of business, reporting behaviour, size, etc.

[45] Another avenue, still to be explored, could be the employment of solutions that allow the visualisation of the connection between the reported data and the discovered anomaly (main variables that apply).

These are only first steps along a path aimed at guaranteeing full transparency and accountability, and further reflection on how these requirements can be enhanced is in our research agenda. More generally, how to automatically improve the explainability of the AI/ML in order to provide RAs with a rather straightforward explanation of the main driver of a potential outlier is a research area that needs further research.

6.3. Technological Enablers: IT Infrastructure and Environment

The availability of appropriate and conducive IT infrastructure and environment to support the testing, deployment and use of the AI/ML solutions must be addressed from different perspectives.

Firstly, AI/ML solutions, especially when particularly complex, demand advanced technological stacks to be developed and then run effectively. Such systems must handle an impressive volume of data and sophisticated algorithms; hence, they must own substantial computing power, storage capacity, and a data catalogue to streamline resource discovery and speed up development.
Secondly, scalability and flexibility of IT platform are other key requirements: they consist in the possibility to access resources on demand, exploit efficient data ingestion tools and, on a more technical perspective, manage libraries autonomously.
Thirdly, modern models like Large Language Models place an additional strain on the infrastructure due to their computational intensity, sometimes making dedicated hardware solutions indispensable.

Lastly, just like any IT application, the AI/ML solution supporting IT system must strictly adhere to the corporate data governance and security policies, for example regarding the handling of confidential information and the protection from external cyber-attacks.

That said, at Banca d'Italia, data scientists can rely on a comprehensive Enterprise Data Lake (EDL)[46]. Its ultimate goal is to allow the combined processing of different data sources, resulting in large volumes and possibly with different structures or, in some cases, even not having a predefined structure at all. The EDL offers all the capabilities required to support modern AI/ML initiatives, ensuring scalability, flexibility, and efficient data management. This facilitates a seamless transition from the prototype and research phase to a more mature stage, where planning new AI/ML solutions for statistical data management becomes a standard activity[47].

Once a new solution has been developed, it is also important to integrate it into the existing statistical workflow so as to avoid it being activated manually, which can negate the time saved by running an automated solution and may burden the data manager with coordination tasks. That being said, we reiterate that the output of the solution should not automatically trigger a further step. Rather it represents evidence for the "human" that evaluates the information content and decides about its follow up. However, based on our experience, the integration of AI/ML solutions into the existing IT infrastructure supporting the statistical processes may be an important challenge, for various reasons**.**

[46] One of the goal of Banca d'Italia's 2023-25 Strategic Plan is indeed to strengthen the IT infrastructure so that it is increasingly suitable to deal with a higher and higher volume and variety of data as well as AI/ML algorithms (Line of action 2.2 "*Create new technological solutions for the integrated exploitation of the Bank's information assets (including company data and their monitoring), in order to improve the analytical capacities of the Bank's various functions, the efficiency of processes and the quality of statistics production, which would also benefit external users*").
[47] The enterprise data lake is available to all the institutional functions of Banca d'Italia, therefore is expected to provide a dramatic efficiency boost also outside the data management perimeter.

Firstly, we have observed that that existing IT solutions - which have been developed before the advent of AI/ML systems - sometimes cannot be easily integrated with new computational AI/ML modules without losing the overall internal consistency of the process. However, a full redesign of the existing solutions is not always desirable as it might be too time consuming and/or costly, particularly when the AI/ML solutions to be integrated are still few. Under these conditions, a viable strategy is to develop an AI/ML solution which fosters its consequent IT implementation "by design".

6.4. The organisational set-up

As far as the organisational set-up is concerned, it is still premature to make a clear statement on whether there is an optimal configuration to adequately support the regular development and adoption of statistical cutting-hedge solutions to enhance the efficiency and effectiveness of the regulatory reporting management in a data-driven institution.

In our initial, "experimental phase", when the development of AI/ML solutions is "on a best effort basis" and ideas are still forming on how and in which areas of activity the new methods can be usefully exploited, it is probably preferable to proceed with a rather unstructured and decentralised approach, i.e. allowing the ideas to form freely and harvest the results; this approach is typical of a "pioneering phase". This is also because, although central banks may use AI/ML to significantly boost their analytical and procedural capabilities, it is also their duty to always rely on sound cost-benefit analyses to ensure that innovations are pursued only as and when they are truly beneficial and sustainable over time. Should a more traditional, simpler and, therefore, presumably less costly solution offer similar results, then it is probably the most suitable way to go (Occam's razor). Moreover, organisational and IT changes in data collection and management processes by the implementation on a vast scale of AI/ML method may put a strain on the company; hence, they can be planned and activated only once sound evidence of their relevance has been accumulated.

If the introduction of AI/ML solutions into operational processes evolves in the direction of becoming an integral part of a company's strategy and culture regarding the collection and management of regulatory data, it is likely that this effort is supported by complementary interventions at organisational level. For instance, it would be worth evaluating the establishment of a dedicated team of experts (data scientists) with the mandate of supporting, overseeing, coordinating and monitoring the implementation of AI/ML solutions in data management processes. This team would cooperate closely with the business units, which represent the actual business cases and ultimately benefit from the new solutions. Moreover, it would act as a catalyst and promoter of new ideas within the business units, leveraging economies of scale and scope in the process. Independently of the actual organisational model that will be adopted, which may vary depending on the specific reference context, when a company commits to moving beyond the experimental stage, in order to fully exploit the potential of AI-ML methods some kind of intervention at the organisational level must be carefully evaluated.

6.5. The benefits

After discussing the hurdles we encountered, we would like to share some thoughts on the benefits that, according to our experience, come from using AI/ML techniques for the management of regulatory reporting: the first two discussed below relate to the effectiveness, the last two to the efficiency of the operational processes.

Firstly, AI/ML models allow for the identification of complex relationships that would otherwise be difficult to detect, by leveraging the contribution of multiple variables or models to generate the output. For example, the use of stacking models for improving the quality of Anacredit data (see par. 2.3) proves to be very effective in detecting subtle anomalies.

Secondly, within the context of DQM, it is feasible to design plausibility check models that are resilient to changes over time in the underlying pattern of data, as they can smoothly handle gradual shifts without the need for abrupt and untimely revisions of DQM thresholds. Moreover, with these models it is possible to define specific thresholds for each RA, hence allowing to take into account, for instance, the peculiarities of their reporting history. This is the case of using Quantile Regression Forests to identify potential outliers in granular banking data (see par. 2.1, par. 2.2).

Thirdly, these techniques often provide the best "quality-to-time" trade-off when dealing with large volumes of data. For example, this is the case of the imputation of the institutional sector (see par. 3.1) as well as of the identification of the nationality of companies (see par. 3.2), both of which would require a tremendous amount of manual work without the support of AI/ML solutions.

Lastly, the possibility to automate manual-intensive decision processes results in time and resource savings which contribute to an improved operational efficiency. For instance, for the time it was active, the project related to the market notices of the Italian stock exchange returned substantial time savings, highlighting the practical benefits of automating previously manual and time-consuming processes (see par. 4.3). When pursuing cost-efficiency management of the operational processes, sparing manual effort and devolving it to activities with a higher "value added" is a primary goal. However, it should be remarked that a labour-saving effect of AI/ML innovation becomes tangible only when a "critical mass" of innovative solutions is implemented, as individually each one might allow to save only a fraction of the human resources involved in the actual operational processes.

7. CONCLUSIONS AND FUTURE DEVELOPMENTS

The paper has surveyed a number of studies completed so far within the statistical function of Banca d'Italia in order to assess the potential of AI/ML methods to enhance the process relating to regulatory data collection from banking and other supervised financial intermediaries. This topic is currently being widely discussed in the statistical community of central banks. In this respect, the paper contributes to this literature not only by providing an overview of the work carried out so far by Banca d'Italia, but also by outlining the main lessons that can be learnt from the actual use cases that have been investigated. Despite the latter refer to a very specific area of potential applications of AI/ML methods, at least some of such lessons are likely to have a more general value.

More specifically:

- firstly, the general approach that was followed to evaluate and implement innovative AI/ML methods can be summarised as "start small, get results, invest in 'enabling factors' and then scale up";

- secondly, the studies carried out so far have been grouped and categorised according to their specific area of research: data validation, enrichment of the register data, and data management efficiency and automation. The main finding of this analysis is that AI/ML can significantly improve efficiency, accuracy, and decision-making in statistical applications relating to regulatory data collection processes;

- thirdly, the main lessons learnt have been summarised. They can be traced back to the fact that the adoption of AI/ML in a central bank data management setting requires a strategic approach in order to set the right pre-conditions for implementing the new methods in operational statistical processes.

  Among the main "enabling factors", the following are especially relevant: promoting a multi-disciplinary approach; investing in the training of specifically skilled staff on an ongoing basis;

guaranteeing the continuous assessment and maintenance of the models, including an evaluation of the implications in terms of turnover of the staff that developed them; guaranteeing model interpretability and transparency, in order to avoid "black-boxes"; taking into account the risks for operational processes that rely on AI/ML solutions due to the variations in characteristics of external data sources, which suddenly might make such solutions "obsolete", putting at risk the continuity of the processes; and investing in IT infrastructure suitable for dealing with AI/ML methods.

Moreover, when designing AI/ML solutions, it is also crucial to keep the IT implementation aspect in the picture, so as to foster their actual integration into the existing operational processes.

In relation to the scale of the challenges and the fact that the business cases often present characteristics common to other central banks, a joint effort between authorities is desirable, since it would provide the opportunity to pull together staff to tackle common challenges and exploit economies of scale.

Finally, it is worth remarking that benefits and costs must be carefully weighed before embarking on any AI/ML project applied to statistical data management, to ensure that innovations are pursued only if they are truly beneficial and sustainable in the long term. The decision should be based on a thorough ex-ante cost-benefit analysis that takes into account the fast-evolving landscape, to ensure that the cost of developing the solution will in fact be outweighed by a tangible increase in efficiency and/or the reallocation of human resources to other valuable tasks. In this regard we observe that the five works[48] (of the nine studies summarised in this note) that have been implemented provide a tangible support to the data management operative processes by reducing the human effort involved and/or increasing the accuracy and completeness of the data. We expect a similar return also from the sixth study[49] which is planned to go-live in the weeks after the publication of this paper.

With regard to future developments of our research program, we will give priority to new data validation algorithms with a view to making data quality procedures more and more precise and, therefore, lighten the burden on RAs. As far as enhancing the efficiency of the data management processes is concerned, a recent and very promising area of research concerns the application of Large Language Models (LLMs). Specifically, we are analysing (i) whether LLMs can be leveraged to reduce the time required in the analysis of the reporting regulation to identify the data quality checks that can be implemented to support the collection phase, and (ii) the implementation of a "Question & Answer" tool that is able to analyse and answer the questions typically raised by RAs in relation to the complex aspects of the regulation concerning supervisory reporting obligations.

Concerning the latter, we have developed a prototype based on a multi-step approach that produces significantly more comprehensive and accurate results than simple "zero-shot" or "few-shots" models with no context enrichment. We are still working to enhance the accuracy of the model to provide analysts with a reliable working tool capable of speeding up the answering process.

[48] Please refer to footnote 16 for the complete list.

[49] Please refere to footnote 16.

REFERENCES

D. Aroujo, G. Bruno, J. Marcucci, R. Schmidt and B. Tissot (2021), Machine learning applications in central banking, *IFC Bulletin*, n. 57.

D. Aroujo, G. Bruno, J. Marcucci, R. Schmidt and B. Tissot (2022), "Data science in central banking: applications and tools", *IFC Bulletin,* n. 59.

J. Brault, M. Haghighi, B. Tissot (2024), "Granular data: new horizons and challenges", IFC Bulletin, n. 61.

Banca d'Italia (2017), Strategic Plan for 2017–2019.

Banca d'Italia (2023), Strategic Plan for 2023- 2025.

M. Bernardini, P. Massaro, F. Pepe and F. Tocco (2021), "The market notices published by the Italian Stock Exchange: a machine learning approach for the selection of the relevant ones", Banca d'Italia, Occasional Papers, n. 632.

A. Carboni and A. Moro (2018), "Imputation techniques for the nationality of foreign shareholders in Italian firms", *IFC Bulletin*, n. 48.

M. Casa, M. Carnevali, S. Giacinti and R. Sabatini (2022), "PUMA cooperation between the Bank of Italy and the intermediaries for the production of statistical, supervisory and resolution reporting", Banca d'Italia, *Occasional Papers*, n. 734.

M. Casa, L. Graziani Palmieri, L. Mellone and F. Monacelli (2022), "The integrated approach adopted by Bank of Italy in the collection and production of credit and financial data", Banca d'Italia, Occasional Papers, n. 667.

R. Caspart, S. Ziegler, A. Weyrauch, H. Obermaier, S. Raffeiner, L.P. Shumacher, J. Scholtyssek, D. Trofimova, M. Nolden, I. Reinartz, F. Isensee, M. Goetz, C. Debus (2022), "Precise energy consumption measurements of heterogeneous artificial intelligence workloads", ISC Workshops 3 Dec 2022.

F. Cusano, G. Marinelli and S. Piermattei (2022), "Learning from revisions: a tool for detecting potential errors in banks' balance sheet statistical reporting", *Quality & Quantity*, January 2022.

Data Management Association (2017), "DAMA-DMBOK: Data Management Body of Knowledge", Second edition, *Technics Publications*.

O. Giudice, P. Massaro and I. Vannini (2020), "Institutional sector classifier, a machine learning approach", Banca d'Italia, *Occasional Papers*, n. 548.

B. La Ganga, P. Cimbali, M. De Leonardis, A. Fiume, L. Meoli and M. Orlandi (2022), "A decision-making rule to detect insufficient data quality: an application of statistical learning techniques to the non-performing loans banking data?", Banca d'Italia, *Occasional Papers*, n. 666.

P. Maddaloni, D. N. Continanza, A. del Monaco, D. Figoli, M. di Lucido, F. Quarta, G. Turturiello (2022), "Stacking machine learning models for anomaly detection: comparing AnaCredit to other banking datasets", Banca d'Italia, Occasional Papers, n. 689.

E. McCaul (2022), “The impact of suptech on European banking supervision”, European Central Bank, Speech at the Supervision Innovators Conference 2022.

E. McCaul (2024), “The future of European banking supervision – connecting people and technology”, European Central Bank, Keynote speech at the Supervision Innovators Conference 2024.

E. Svezia and V. La Serra (2024), “A supervised record linkage approach for anomaly detection in insurance assets granular data”, *Quality & Quantity*, March 2024.

United Nations Economic Commission for Europe (2022), *Machine Learning for Official Statistics*, UNECE, New York and Geneva.

United Nations Economic Commission for Europe (2023), *Large Language Models for Official Statistcs*, UNECE, New York and Geneva.

European Union (2024), "Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonized rules on Artificial Intelligence (AI Act)", Official Journal of the European Union, L 257.

F. Zambuto, S. Arcuti, R. Sabatini and D. Zambuto (2021), “Application of classification algorithms for the assessment of confirmation to quality remarks”, Banca d’Italia, Occasional Papers, n. 631.

F. Zambuto, M. R. Buzzi, G. Costanzo, M. di Lucido, B. La Ganga, P. Maddaloni, F. Papale and E. Svezia (2020), “Quality checks on granular banking data: an experimental approach based on machine learning”, Banca d’Italia, Occasional Papers, n. 547.