

# ENHANCING TAX AUDITS: PREDICTING TAXPAYER BEHAVIOUR THROUGH AI

Speaker: *Annalisa Mellace*

Italian Revenue Agency  
Data Science Office



# OUTLINE

- The main goal
- The data
- The solvency score
- The model
- National - wide distribution
- Performances and value added
- Future developments

# THE MAIN GOAL

How do we optimize resources to ensure the greatest impact?

- 1** Provide a reliable forecast of the possible behaviour that taxpayers might adopt in the event they receive a formal notification of a tax assessment, or if a direct investigative activity is initiated against them.
- 2** Provide highly targeted operational assistance in determining the exact processing priority for a vast number of cases concerning VAT-registered entities that are formally included in our **Annual Audit Plans**.

**3**

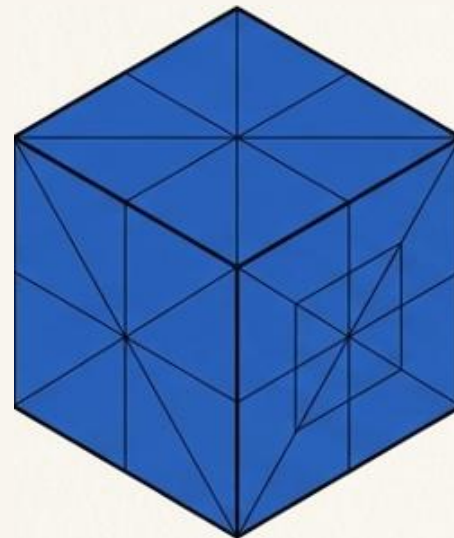


# DATA COLLECTION

- Tax Returns
- E- Invoices
- Asset data
- Payment Reminders
- 
- 
- 
- Behaviour in past assessment

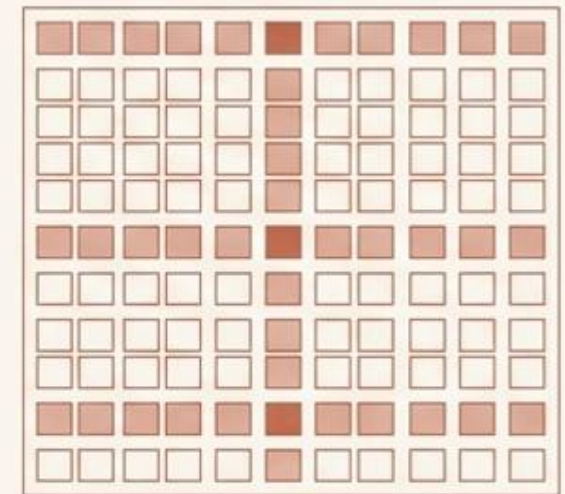
INPUT DATA

ETL WORKFLOW



DATASET

Distilling raw interaction into ~300 specific behavioural and economic variables per taxpayer.



# THE WORKFLOW



## STEP 1: ETL

*SPSS Modeler*

Extracting, transforming and loading vast amounts of historical data.



## STEP 2: PREDICTIVE TRAINING

*Python*

Predictive models trained to identify complex, non-linear relationships.



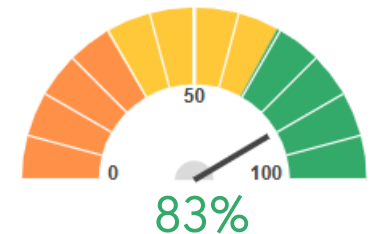
## STEP 3: ALGORITHMIC ENGINE

Execution of the algorithm and prediction of the position included in the Audit Plans.

## FINAL STEP: PERFORMANCE

Prediction on unseen taxpayer positions to validate the consistency of the models.

*Predictive Accuracy:*



# THE DYNAMIC SOLVENCY SCORE

$$Y = \frac{\text{Collected payment}}{\text{Requested amount}}$$

**Issue:** During preliminary evaluations, the final assessment amount is unknown.

**Solution:** the algorithm adapts by generating predictive scores across seven theoretical financial demands.

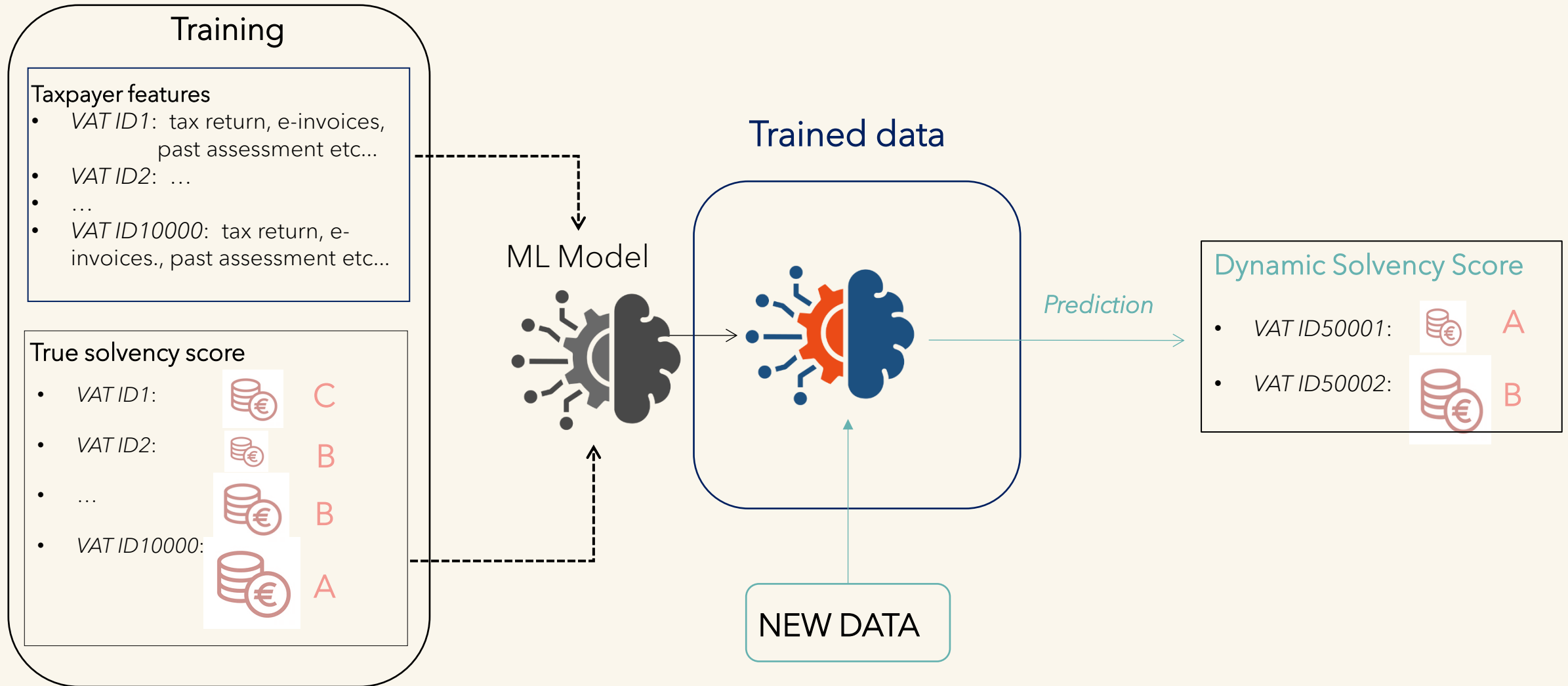
## 3 POSSIBLE OUTPUTS

- $\hat{Y} =$
- A: High probability of full payment
  - B: Partial payment expected
  - C: No expected payment

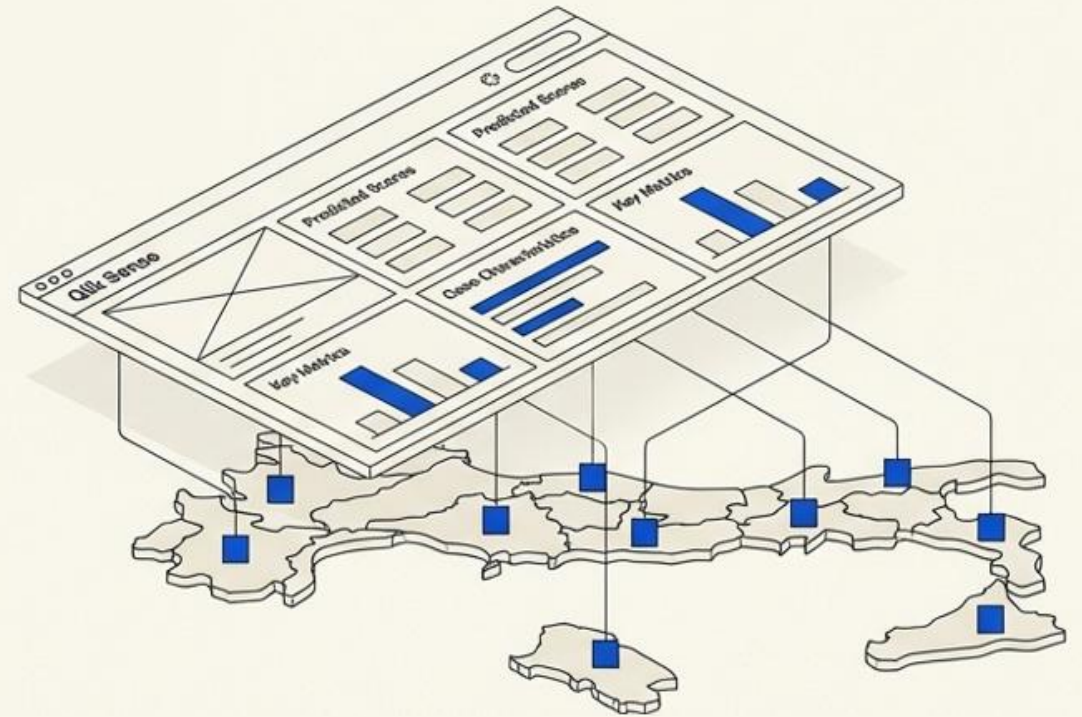
## 7 CLASS OF REQUESTED AMOUNTS

<€1k	€1k-€5k	€5k-€20k	€20k-€50k	€50k-€250k	€250k-€500k	>€500k
------	---------	----------	-----------	------------	-------------	--------

# THE MODEL



# INTERACTIVE DASHBOARD FOR A NATION-WIDE USE



## The Interface

Built on Qlik Sense for business intelligence and exploration. Officers can view the predicted scores alongside key case characteristics.

## The updates

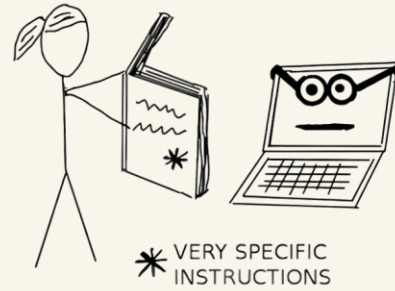
System updates automatically every two weeks to capture the latest fiscal events.

## The Expansion

Expanded in January 2026 to include taxpayers without a VAT number, covering almost the entire taxpayer base.

# THE SHIFT

Without Machine Learning



With Machine Learning



## 3 PROS

Analytical Scale & Processing Time



### TRADITIONAL APPROACH

Manually process limited amount of data for each single taxpayer.

Adaptability & Continuous Learning

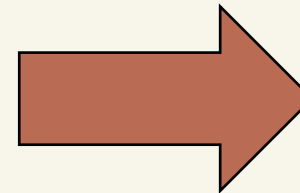


Slow adaption to rapidly changing economic scenarios and taxpayer behaviours.

Consistency & Decision Support



Difficult to maintain a perfectly uniform baseline across all regions.



### AI APPROACH

The model instantly analyzes ~300 variables across the entire country.

ML rapidly adapts to new data, organically improving its predictive accuracy over time.

It guarantees structural consistency, allowing officers to focus their valuable professional experience on the final decision.

# RESULTS

March 2025 - March 2026

## User Adoption

- Feedbacks from the final users:

*«The tool is actively used and considered highly effective for accelerating and simplifying the prioritization of the Annual Audit Plan.»*

## Proven Predictive Accuracy

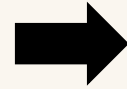
- Real case validation: after one year, initial AI predictions were compared against the actual outcomes of the completed audits.
- Performance: Over **82%** of the model's predictions proved to be correct in the field.

High satisfaction with the system's reliability as a concrete decision-support tool.

# FUTURE DEVELOPMENT

## Monitoring

- ❑ Observe the algorithm recently extended to include individual taxpayers;
- ❑ Track model performance, monitor field usage and gather user feedback.



## Validation

- ❑ Conduct a comprehensive 12-month validation study by early 2027 on future real data;
- ❑ Gather final user feedback.



## New features and integration

Integrating the solvency analysis of a corporate entity with the fiscal behaviour of its shareholders and partners.

# THANK YOU!

*Annalisa Mellace*

For any question, please feel free to reach us here:  
*annalisa.mellace@agenziaentrate.it*

Italian Revenue Agency  
Data Science Office

