Plenary Session 5: Blockchain, Artificial Intelligence, and Machine Learning in Tax Services and Enforcement

The integration of Blockchain, Artificial Intelligence (AI), and Machine Learning (ML) in Taxpayer Services and Compliance is a topic of great significance in today's rapidly evolving technological landscape. This session aims to provide a comprehensive overview of how these innovative technologies can enhance the process of tax compliance and enforcement. It will cover the role of blockchain, AI, and ML in shaping the future of tax compliance and enforcement including their potential in areas of detection of non-compliance, taxpayer education and efficiency in tax administration.

Presentation

- Mr. Reriton Weldert Gomes, General Coordinator – Registry and Tax Benefits, Federal Revenue of Brasil

- Mr. Giovanni Tencio, Director – Tax Audit, Directorate General of Taxation (DGI), Costa Rica

Panel discussion

- Mr.Datuk Mohd Nizom Sairi, Chief Executive Officer/Director General, Inland Revenue of Malaysia

- Ms. Nena Dokuzov, Head – Blockchain Technology Group, Ministry of Economic Development and Technology, Slovenia

Open discussion and Q&A

Moderator

- Ms. Melinda Crane, Moderator, Deutsche Welle

Presentations

Tencio, G. (2023). *Using machine learning to manage compliance risk: the case of the electronic invoice anomaly detector: the Costa Rica experience*. Ministerio de Hacienda Costa Rica.

Weldert Gomes, R. (2023). *Blockchain applied to the taxpayer registry*. Federal Revenue of Brazil.

2nd nto Technical Conference
hosted by ATAF

DIGITALISATION OF TAX ADMINISTRATIONS AND CONTEMPORARY ISSUES

05-07 September 2023

Blockchain applied to the tax payer registry

RERITON WELDERT GOMES

General Coordinator of registry and tax benefits

Federal Revenue of Brazil

Funded by

Facilitated by

Hosted by

Partners

Blockchain benefits for registry

Simplicity

Reduced economic costs, time and complexity in inter-governmental and public private exchange registry information

Efficiency

Less bureaucracy in sharing information with other entities

Trust

Automation, transparency, and accountability of governmental registries for the benefit of citizens

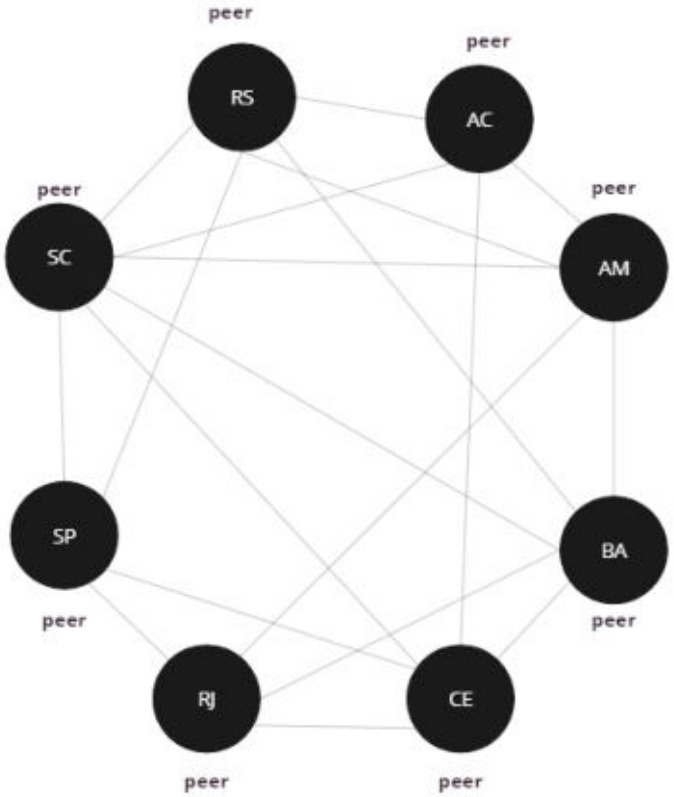
Security

Security in the registry exchange process, improving public services and the business environment

Integrated and cooperative solution

b-Cadastros (CIN) - permissioned blockchain with a peer for each state identification institute in brazil

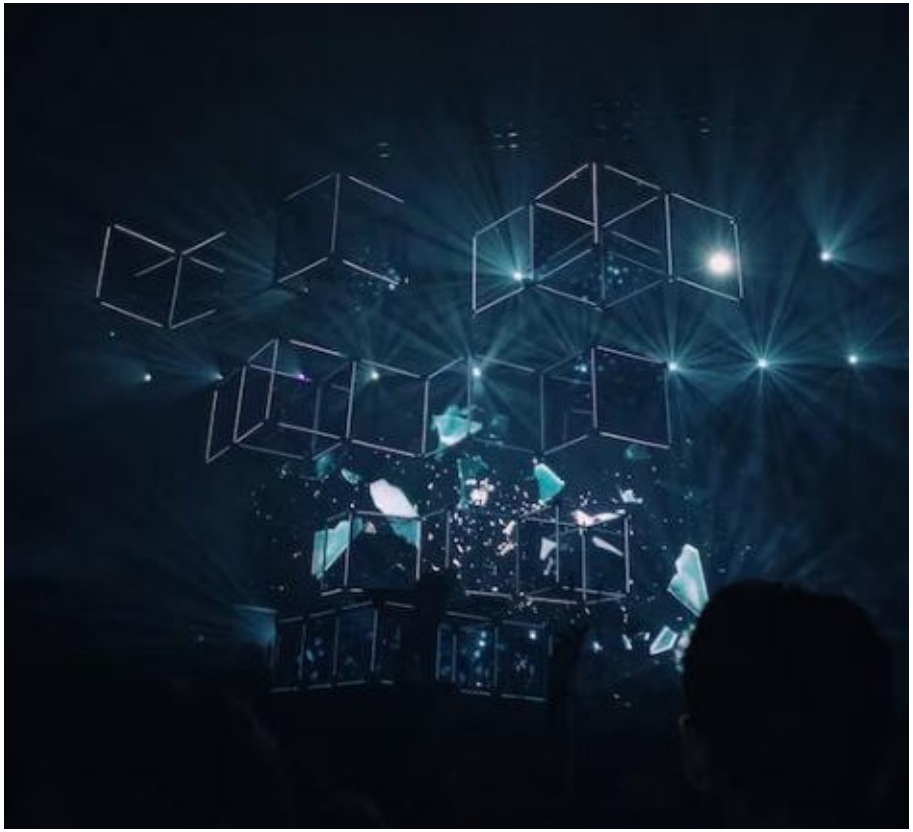
Each of the 27 peers, communicates with the other nodes through the b-Cadastros network and is always up-to-date on new issues and changes to CINs or registrations and changes to CPFs carried out by the other OICs.



B-CADATROS

- Secure data sharing for the issuance of the brazilian *'National Identity Card'* (CIN – Carteira de Identidade Nacional);

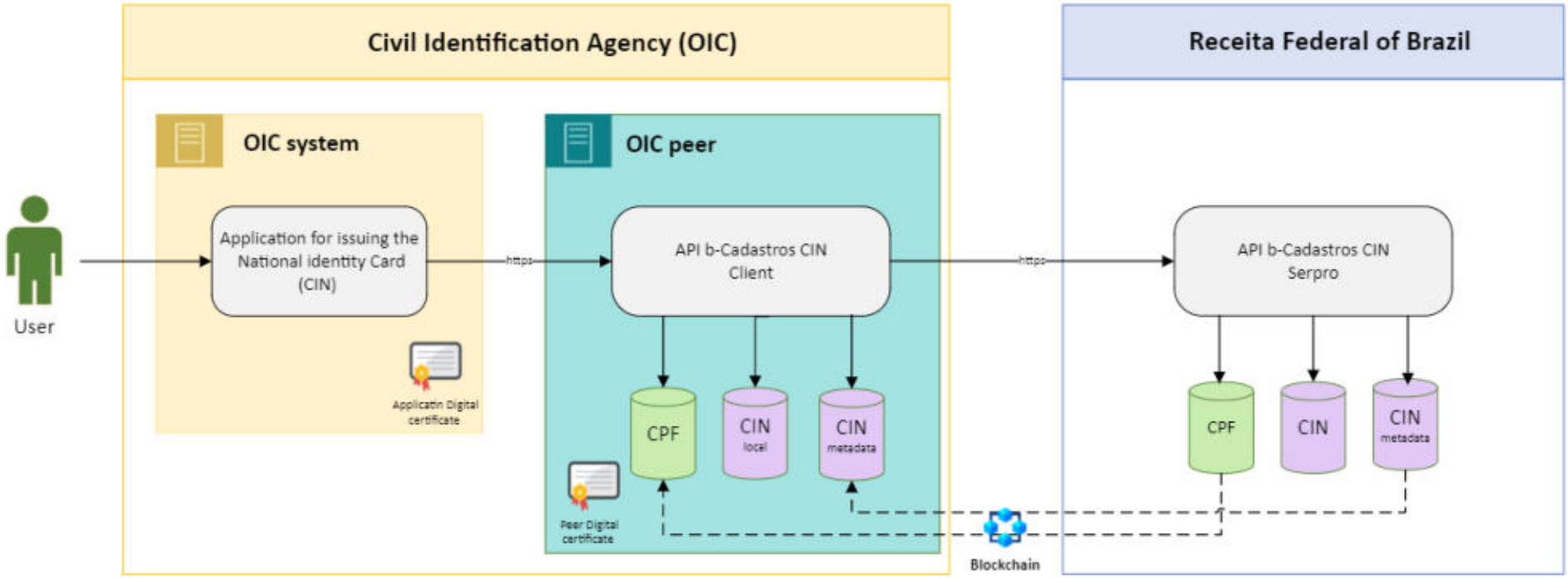
- Innovative, blockchain-based system for a unified and decentralized process.



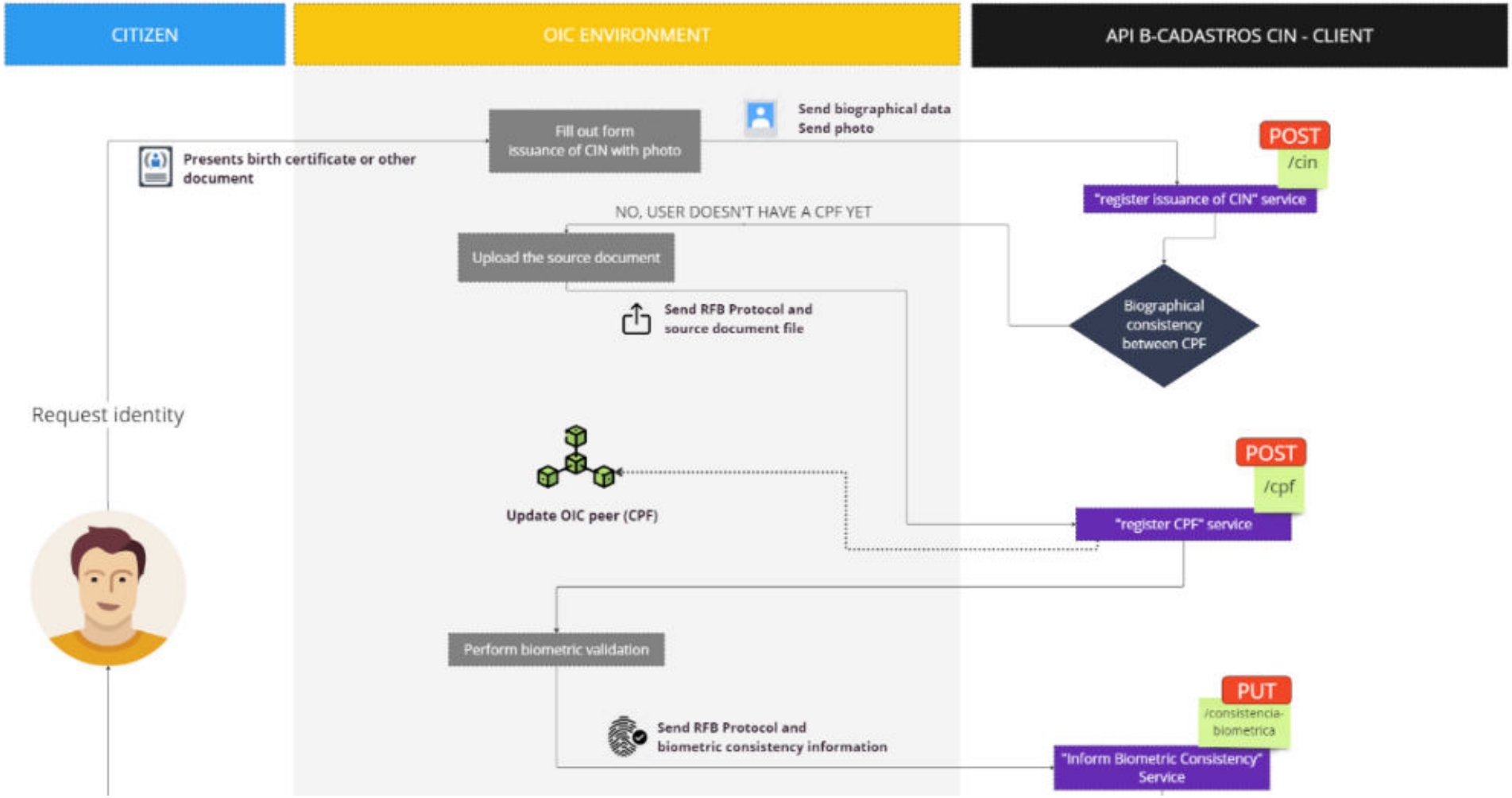
B-CADASTROS

- This solution provides secure data sharing between the RFB and the Civil Identification Agencies (OICs), allowing the National Identity Card (CIN) issuance process to be unified and standardized across all agencies in Brazil.

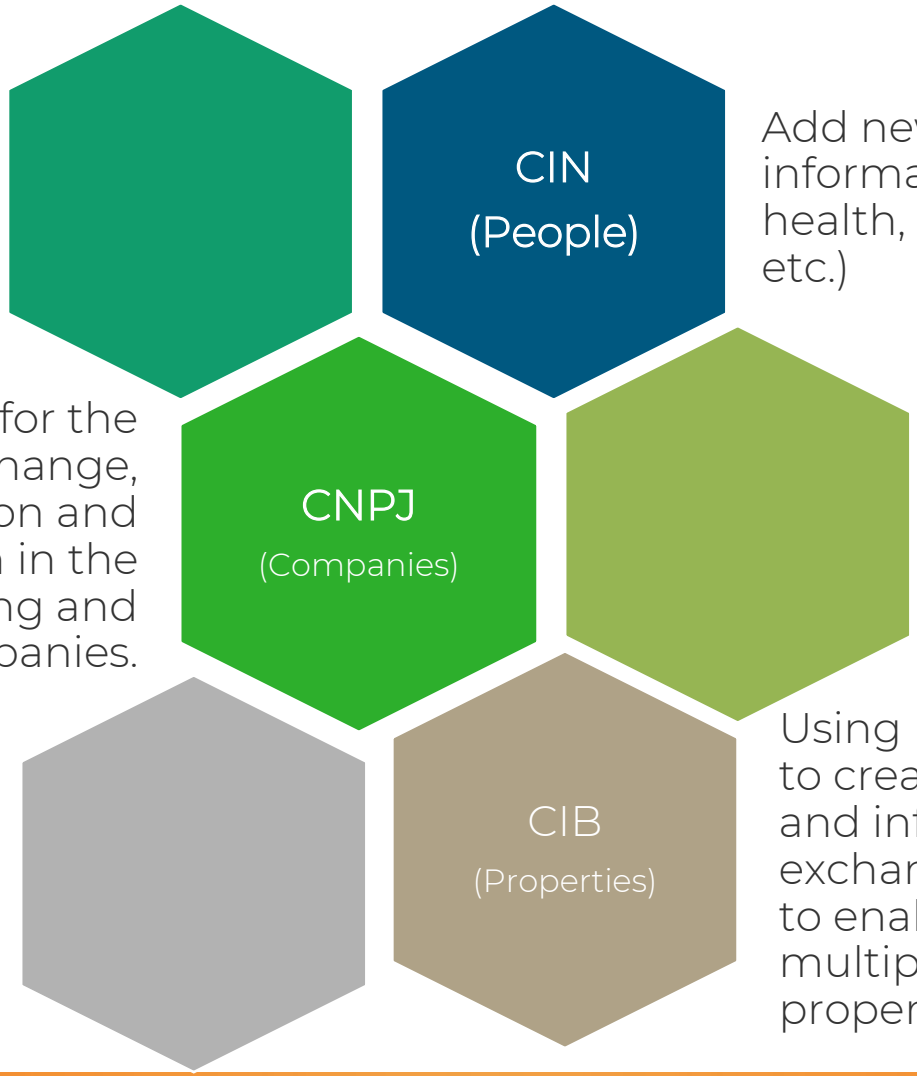
Solution‘s software architecture



CIN Issuance Business Flow



Next use and applications



Add new attributes and information (education, health, social benefit, etc.)

Use the Blockchain for the collection, exchange, alteration and communication in the process of opening and legalizing companies.

Using Blockchain to create a data and information exchange network to enable a multipurpose property registry

2nd nto Technical Conference hosted by ATAF

Thank you

Receita Federal

Funded by

german cooperation
DEUTSCHE ZUSAMMENARBEIT

Co-funded by the European Union

Facilitated by

itc

Hosted by

AFRICAN TAX ADMINISTRATION FORUM
FORUM SUR L'ADMINISTRATION FISCALE AFRICAINE

Partners

A·T·A·I·C

cata

CIAT

COTA

CREDAF

IOTA
Intra-European Organisation of Tax Administrations

PITAA

WATAF FAFOA

2nd nto Technical Conference hosted by ATAF

2nd NTO Technical Conference

Funded by

german cooperation
DEUTSCHE ZUSAMMENARBEIT

Co-funded by the European Union

Facilitated by

itc

Hosted by

AFRICAN TAX ADMINISTRATION FORUM
FORUM SUR L'ADMINISTRATION FISCALE AFRICAINE

Partners

A·T·A·I·C

ata

CIAT

COTA

CREDAF

IOTA
Intra-European Organisation of Tax Administrations

PITAA

WATAF
PAFOA

2nd nto Technical Conference hosted by ATAF

Using machine learning to manage compliance risk: the case of the electronic invoice anomaly detector

The Costa Rican experience

Ministerio de Hacienda COSTA RICA

Funded by
german cooperation DEUTSCHE ZUSAMMENARBEIT
Co-funded by the European Union
Facilitated by
itc
Hosted by
AFRICAN TAX ADMINISTRATION FORUM
FORUM SUR L'ADMINISTRATION FISCALE AFRICAINE
Partners
A·T·A·I·C
ata
CIAT
COTA
CREDAF
IOTA Intra-European Organisation of Tax Administrations
PITAA
WATAF PAFOA

Use of apocryphal invoices

- Problem present in all tax administrations.
- Illegal and hidden activity
- Present in all taxpayer strata.
- Tax simplification and online services facilitate the registration of obligors and the issuance of invoices.
- The electronic invoice avoids the alteration of invoices and the theft of identity, but not the issuance of invoices that respond to false or simulated operations.
- Evolution towards networks for issuing false invoices.

Electronic Invoicing Anomaly Detector

Argentina
España
Brasil
Guatemala
Chile
México
Costa Rica

Open-source solution for CIAT members

Using the e-IAD

- Detection of opportunities for improvement in the quality and structure of the data.
- Validations and alerts, design of outreach campaigns, input for developers,
- Identifies taxpayers with anomalous behavior. Improvement in the selection of audit cases.
- No previous knowledge is required - unsupervised process,
- Understanding of taxpayers and their business cycles.
- Developing analytical skills. Few fixed costs.

Definition

This solution makes it possible to identify, select, and prioritize cases of taxpayers with unusual behavior, by applying unsupervised automatic learning models to data from electronic invoices, taxpayer registration, and tax returns. Although the model is not trained to search for a certain characteristic (unsupervised model), indicators and variables are calculated to search for anomalous behavior in them.

Examples of anomalies that may be detected

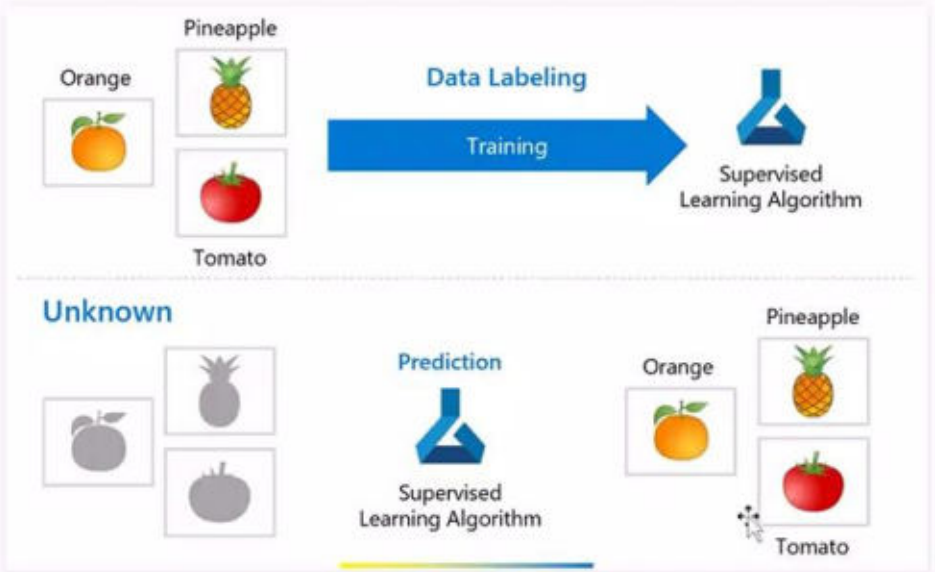
1. Taxpayers with non-existent or simulated operations (suspicious invoicing)
2. Taxpayers receiving income for unusual behavior

Statistical view of loaded data

The Costa Rican case: 1.12 billion records processed in less than seven hours (thirty-three months of data)

Supervised or Unsupervised ML approach for e-Invoicing data

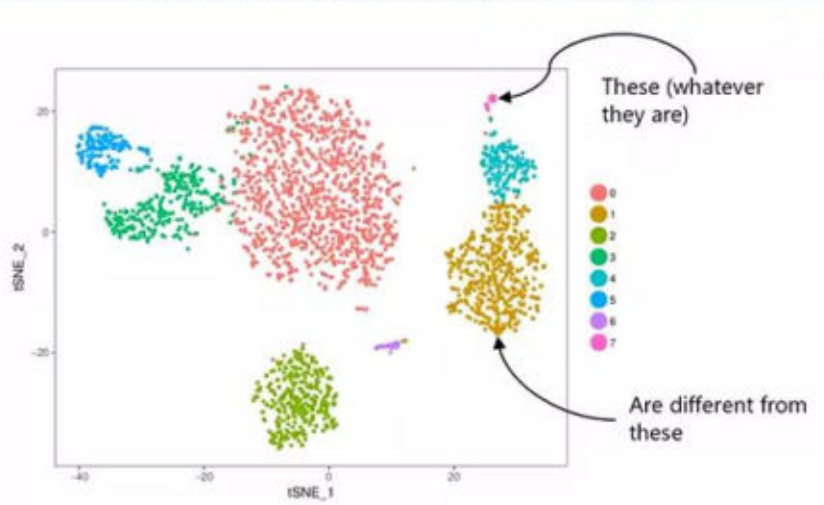
Companies that run intentional or unintentional abnormal transactions emulate regular business behavior – Then, how do you define risk conditions?



Supervised Machine Learning

The algorithm learns from known examples.
Do we have enough labeled data to train the model?

e-IAD introduces an unsupervised model



Unsupervised Machine Learning

The algorithm learns to identify patterns without prior knowledge. *How to improve accuracy?, How to facilitate human intervention?*

E-Invoicing feature engineering spec.

Summarization Features

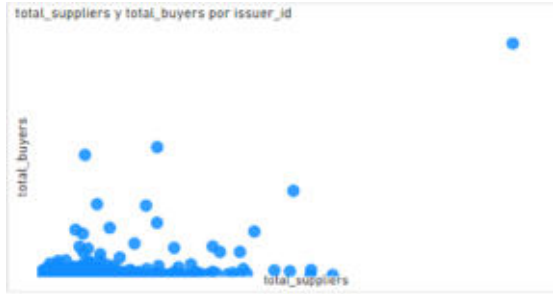
- Number of transactions
- Total invoices
- Total credit notes
- Total debit notes
- Total export invoices
- Total purchase invoices
- Total buyers
- Total suppliers
- Number of transactions to self
- Total voucher to self

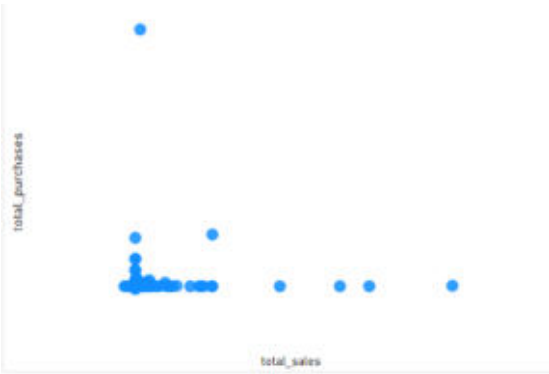
Taxable Features

- Total taxable services
- Total taxable goods
- Total tax
- Total non taxable services
- Total non taxable goods
- Total taxable
- Total non taxable
- Total sales
- Total discounts
- Total voucher

Ratio Features

- Number of purchases
- Total purchases
- Ratio sales / purchases
- Ratio buyers / suppliers
- Ratio VAT / sales
- Ratio sales / employees
- Ratio sales / capital





E-Invoicing feature engineering spec.

Average features

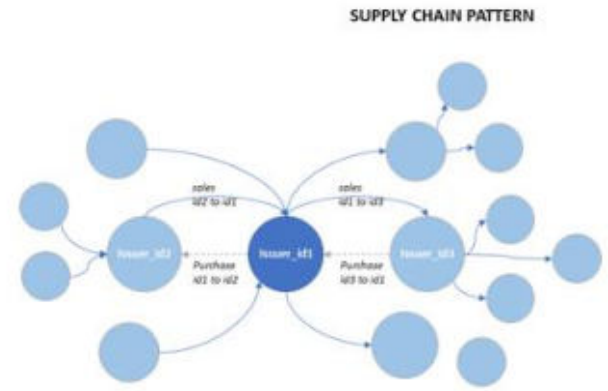
- Monthly total voucher
- Monthly total purchase
- Monthly total tax
- Average monthly voucher
- Average daily voucher
- Average monthly tax
- Average daily tax

Emission features

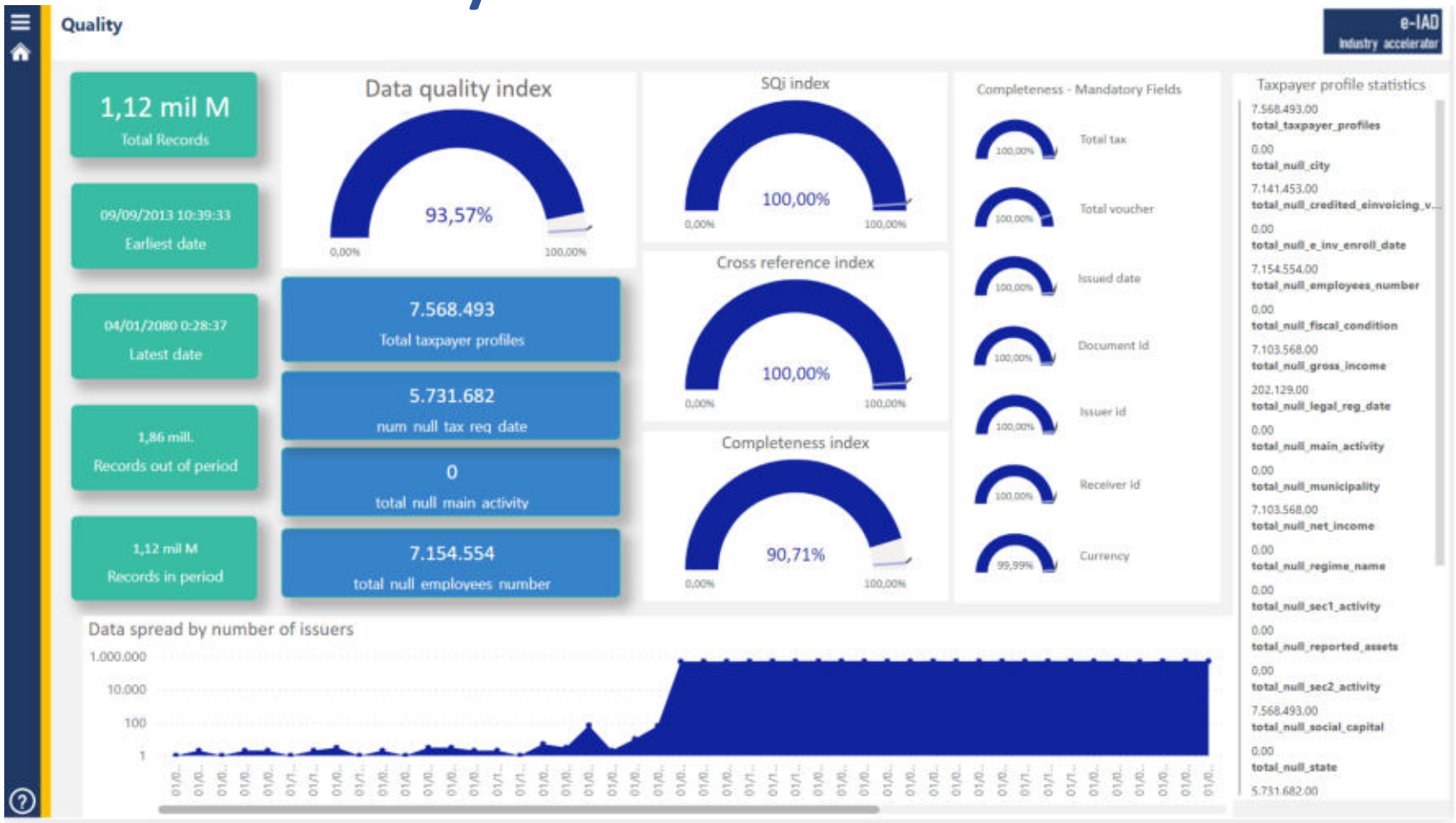
- First issued date ever
- First issued date period
- Last issued date period
- Last issued date ever
- Num days issuing ever
- Num days issuing period
- Time to issue reg
- Time to issue enroll

Graph Features

- Max depth of supply chain
- Max distance from supplier
- Max distance from customer
- Max place in supply chain
- Min depth of supply chain
- Min distance from customer
- Min distance from supplier
- Min place in supply chain

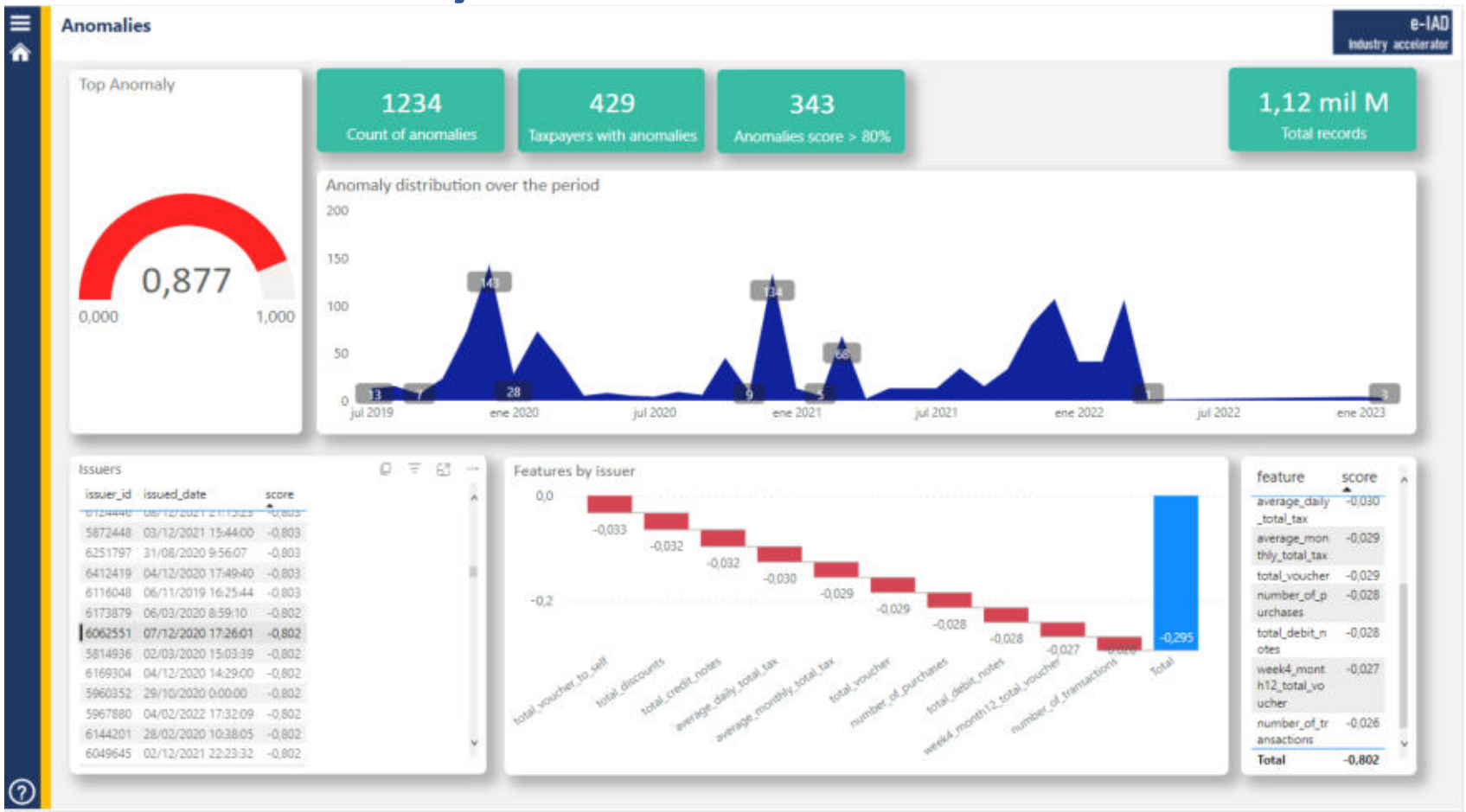


Functionality



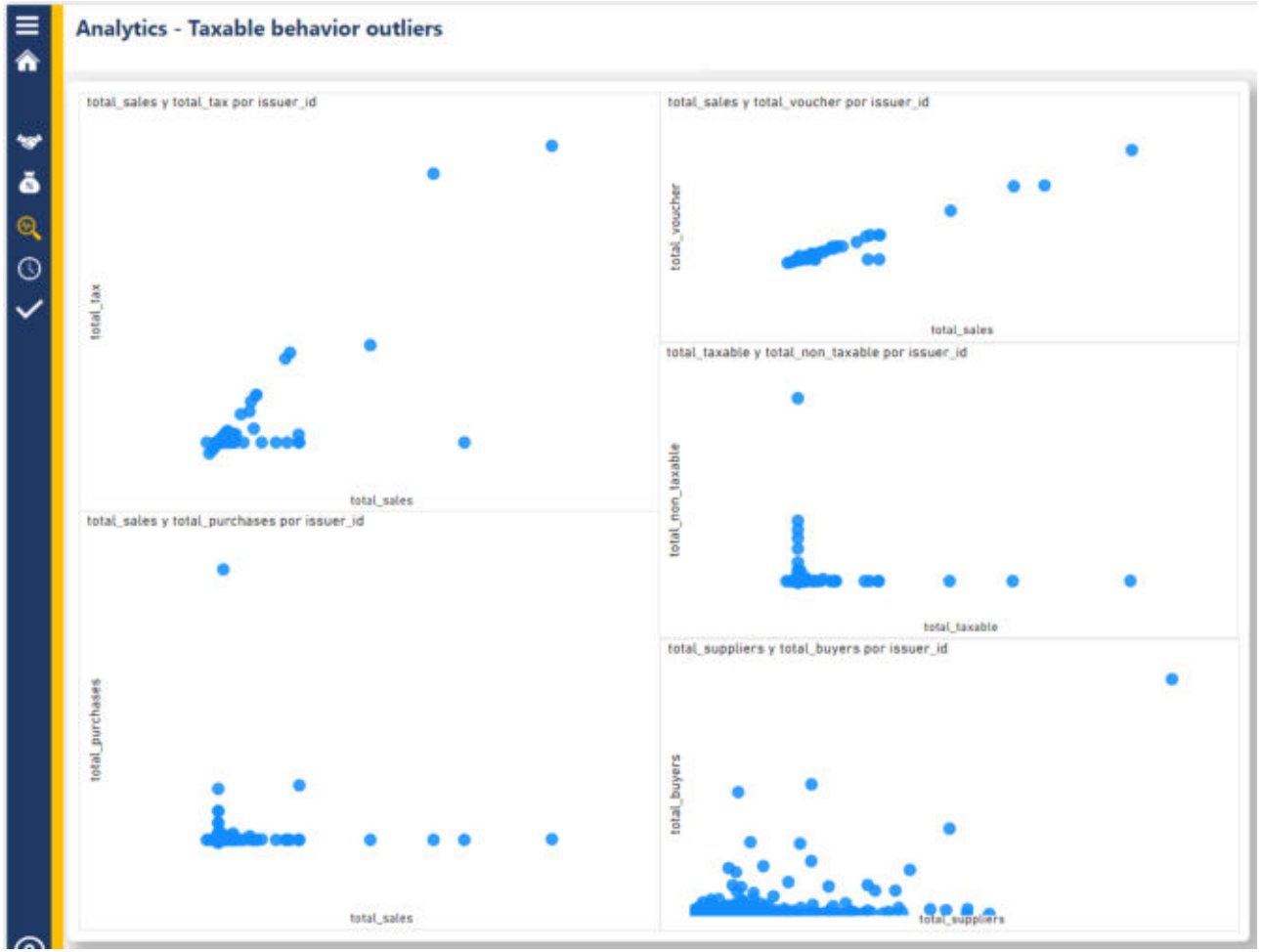
The quality of the data is evaluated from 3 perspectives: completeness, identification, in/out range

Functionality



The system prioritizes cases through an anomaly index, similar to a risk ranking.

Functionality

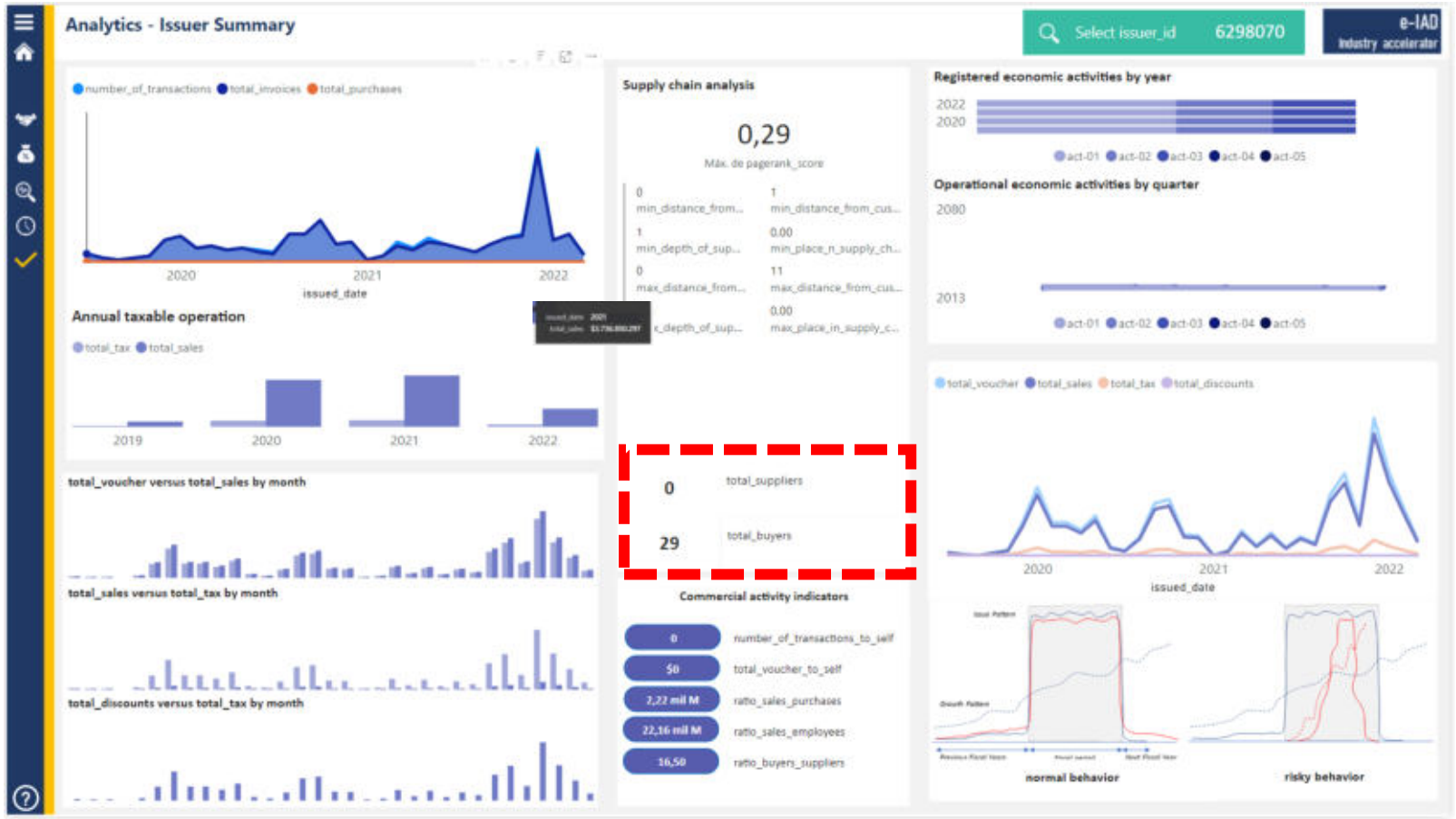


- Prioritize the analysis to taxpayers with extreme values.
- 5 predesigned charts to analyze the relationships between sales, VAT, purchases, number of customers and suppliers
- Filters by state, municipality and city, by economic activity and by year

Findings

Example of a taxpayer with an abnormal behavior of their sales, and issuance of credit notes

Findings

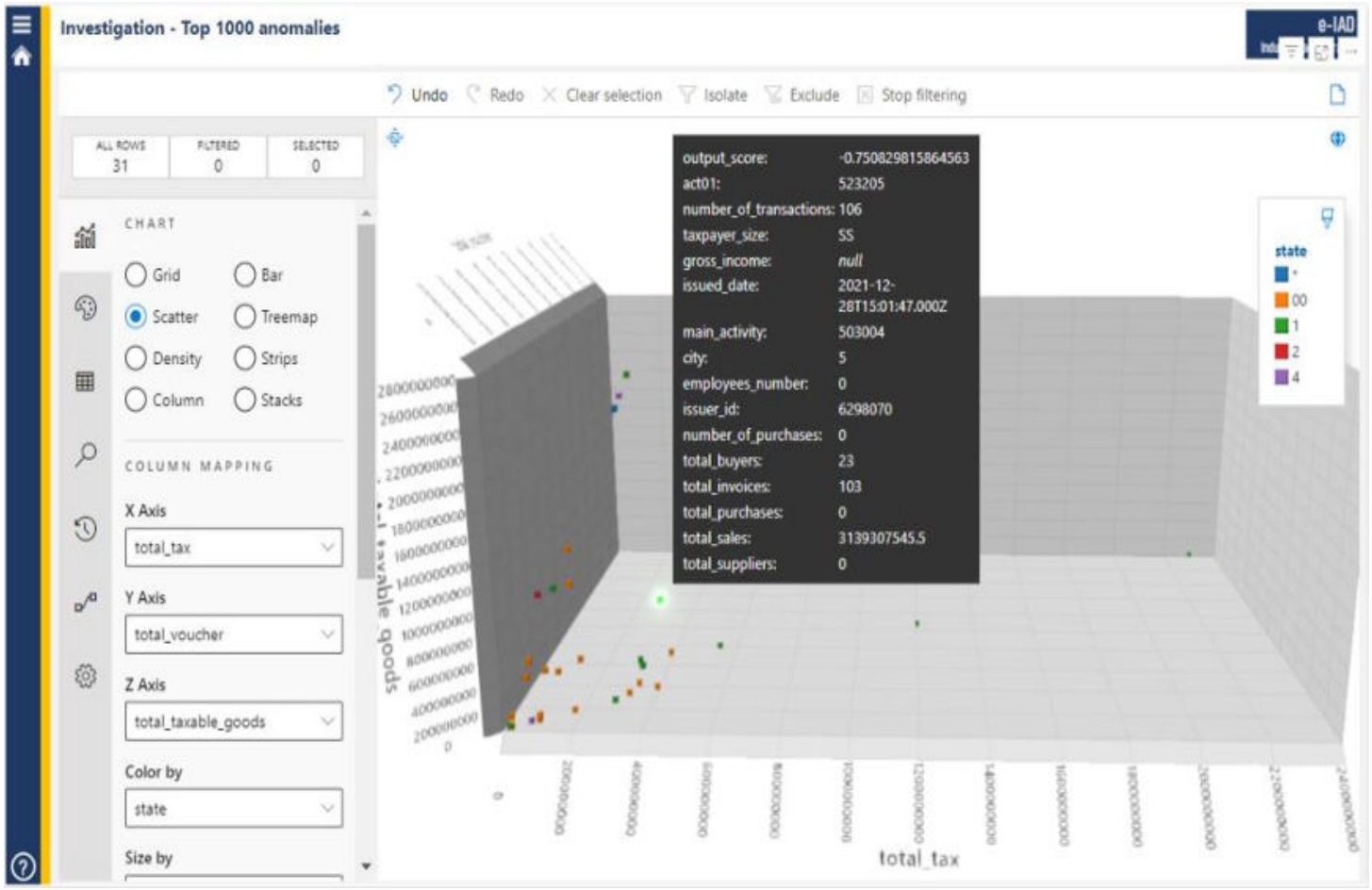


The prioritized taxpayers are analyzed through descriptive statistics and graph theory (study of relationships between entities).

Example of the detection of a taxpayer with **anomalies in his supply chain** (no suppliers and 29 buyers).

Upon further analysis, it is determined that it has no employees, assets or capital

Findings



Issues vouchers (sales) for :

USD$ 5,782.605

- Does not register purchases
- Has no employees

Findings

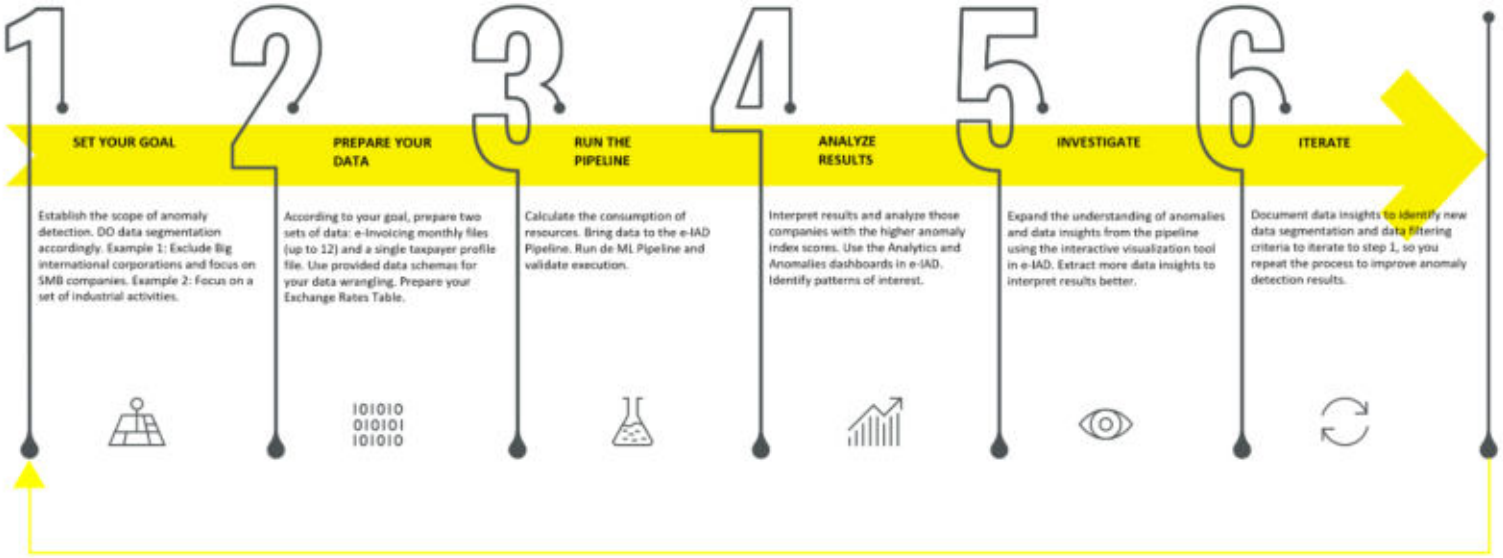
- Companies with irregular behavior in their sales activities that are not sustained by their compliance activity.
- Companies with an atypical supply network, where the volume of customers and /or suppliers is out of proportion.
- Companies with an atypical volume of suppliers.
- Companies with an atypical volume of buyers.
- Companies with an atypical volume in the issuance of credit notes or debit notes.
- Companies with an atypical volume of discounts.
- Companies with atypical amounts in their tax rate.
- Companies with atypical amounts in their services or non-taxable products.
- Companies with sales operation volumes outside the “company size" range in which they are typified.
- Companies with an atypical volume of invoices issued to themselves.

Ministerio de Hacienda COSTA RICA

2nd nto Technical Conference

ted by ATAF

Help



Why

The Microsoft e-Invoicing Anomaly Detection Solution Accelerator will help Tax Agencies in the Civilian Government that has implemented electronic invoicing to detect invoicing anomalies and check large data sets for consistency to help improve revenue collection. The main challenges for effective adoption of AI/ML on e-Invoicing data are the ability to process large datasets at scale and reduce non-time series data's multidimensionality. The well-known risk model approach to selecting tax documents or taxpayers for analysis and investigation also suffers from the challenges of big sets of data.

How

The e-IAD aims to address these typical challenges and, most importantly, enable an automated ML pipeline. This pipeline will facilitate the data loading, calculate the data quality, enrich, and expand the data features to do ML, and then apply an ML algorithm for anomaly detection and produce meaningful results that support the tax agency investigation process. A comprehensive set of visualizations will help with the interpretation of results and the analysis of data along the entire pipeline. These functions enable revenue agents to recognize patterns that could indicate abnormalities in electronic invoices or similar data sources and then flag those issues for investigation.

What

Primary use cases:

- Reduce data processing complexity for analytics to detect anomalie problems faster.
- Measure e-Invoicing data quality with a rapid e-Invoicing analysis tl agencies identify policy gaps in collecting e-Invoicing documents.
- Represent B2B, B2C, and B2G commercial transactions in the econc agencies to describe and analyze issuing patterns and commercial t
- Detect e-Invoicing anomalies in the tax agency's interest to identify

DIGITALISATION OF TAX ADMINISTRATIONS AND CONTEMPORARY ISSUES

2nd nto Technical Conference hosted by ATAF

Funded by

german cooperation
DEUTSCHE ZUSAMMENARBEIT

Co-funded by the European Union

Facilitated by

itc

Hosted by

AFRICAN TAX ADMINISTRATION FORUM
FORUM SUR L'ADMINISTRATION FISCALE AFRICAINE

Partners

A·T·A·I·C

ata

CIAT

COTA

CREDAF

IOTA
Intra-European Organisation of Tax Administrations

PITAA

WATAF
PAFOA