

**TRIBUNAL DE CONTAS DA UNIÃO**  
**Secretaria-Geral de Controle Externo**  
**Secretaria de Controle Externo da Previdência, do**  
**Trabalho e da Assistência Social**

**Continuous Supervision on Social Benefits**



# Structure of the presentation

- **Context**
  - What is TCU?
- **Methodology:**
  - Government Auditing Generations
- **Cycle 2015 – Results:**
- **Cycle 2016:**
  - Pensions
  - Social Assistance
  - Labor



# Context TCU

President's Office



Ministries



Congress



TCU (Supreme Audit Institution)



Supreme Court



# Context

## Annual Budget & Numbers

- **Pensions**

- USD 138 bi/yr
- +30 mi pymt/mo

- **Labor**

- USD 16.7 bi
- +600 k pymt/2wk

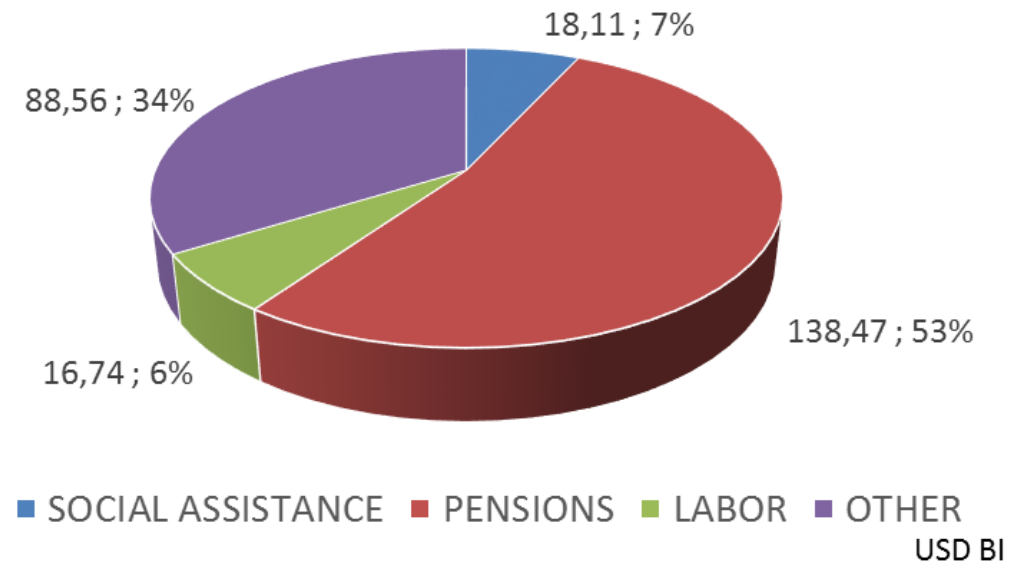
- **Social Assistance**

- USD 18.1 bi
- CadUnico: 120 mi
- Bolsa Familia: 25 mi pymt/mo

- **Debt related expenditure: USD 328 bi**

- **Total: USD 590 bi**

GOVERNMENT BUDGET 2015  
(W/O DEBT RELATED)



# Methodology

## Previous works

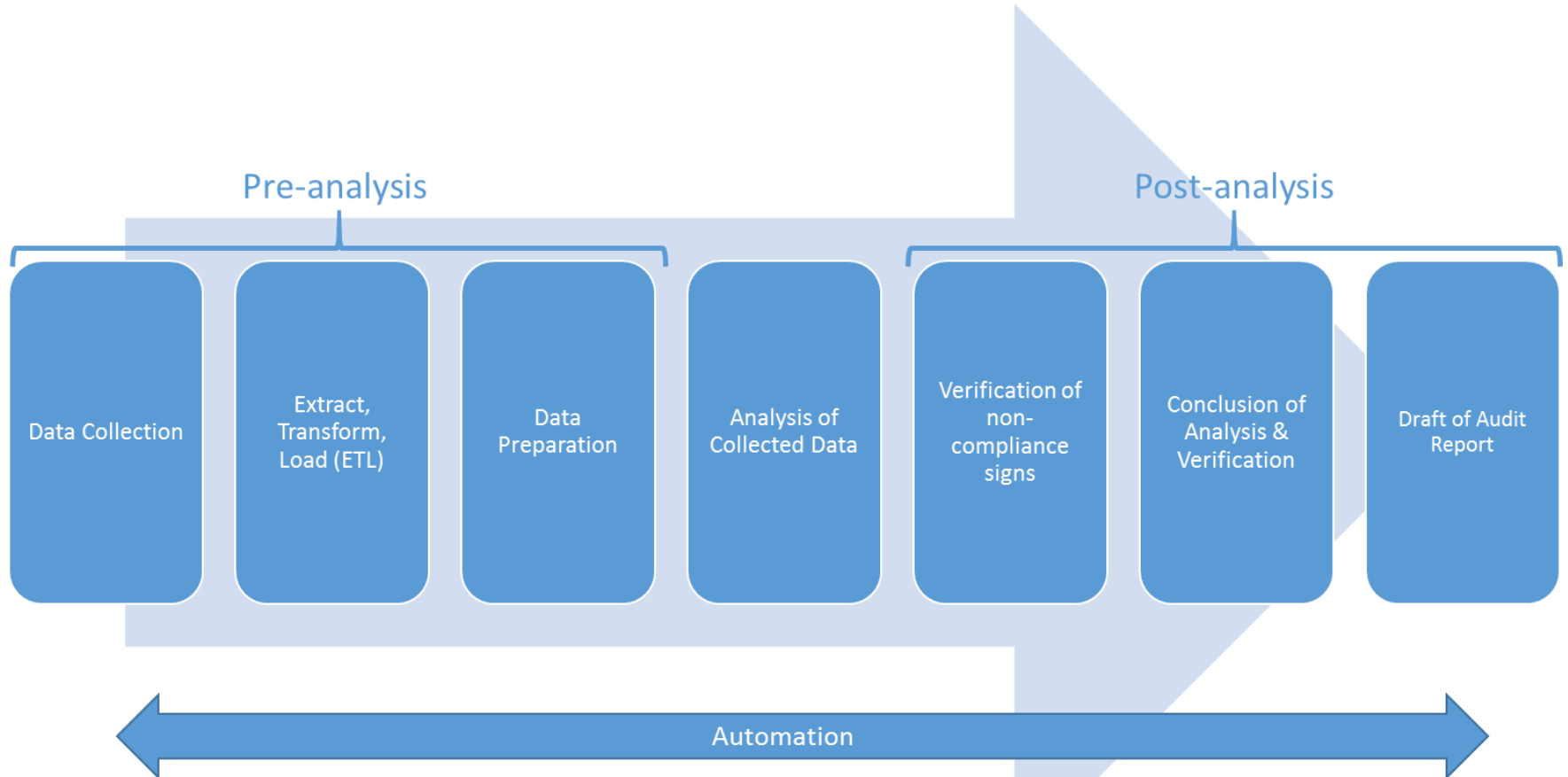
- Continuous Auditing x Continuous Supervision
- Types of Government Auditing: Compliance, Performance, Financial
- Evolution of techniques in Governmental Data Auditing of Social Benefits:
  - 1<sup>st</sup> Generation (until 2006)
    - Traditional auditing of physical documents
  - 2<sup>nd</sup> Generation (2006-2015)
    - Moderate use of IT tools (crossing of few databases)
  - 3<sup>rd</sup> Generation (2015-2016)
    - Automatization of procedures (increase in volume of data)
  - 4<sup>th</sup> Generation (2016-)
    - Use of statistical and machine learning techniques (increase in volume and complexity of data and new expected products)

# Methodology

## Continuous Supervision

- Credibility
  - Public databases show high rate of missing values and contradictory registration data
  - Importance of Data Quality techniques and Credibility indicators
- Timing of auditing
  - Massive amount of payments (pensions and social assistance benefits) and short cycles of benefit granting (unemployment benefits)
  - Few audits per year vs. Continuous monthly (weekly) checks
- Fraud complexity
  - Use of forged documents
  - Organized groups of fraudsters
  - Patterns of fraud

# Methodology Fluxogram



# Cycle 2015

## Results – Pensions

- Implementation of 4 non-compliance typologies
  - Undue accumulation of benefits
  - Benefit with NIT (worker ID number) shared by more than one holder
  - Benefit value over maximum allowed
  - Rural benefit unduly payed to urban worker
- Possible yearly savings of USD 48 mi
- Moderate use of Data Quality and Automation tools
- Implementation of a Pension Panel (BI tool for auditing)



# Cycle 2015

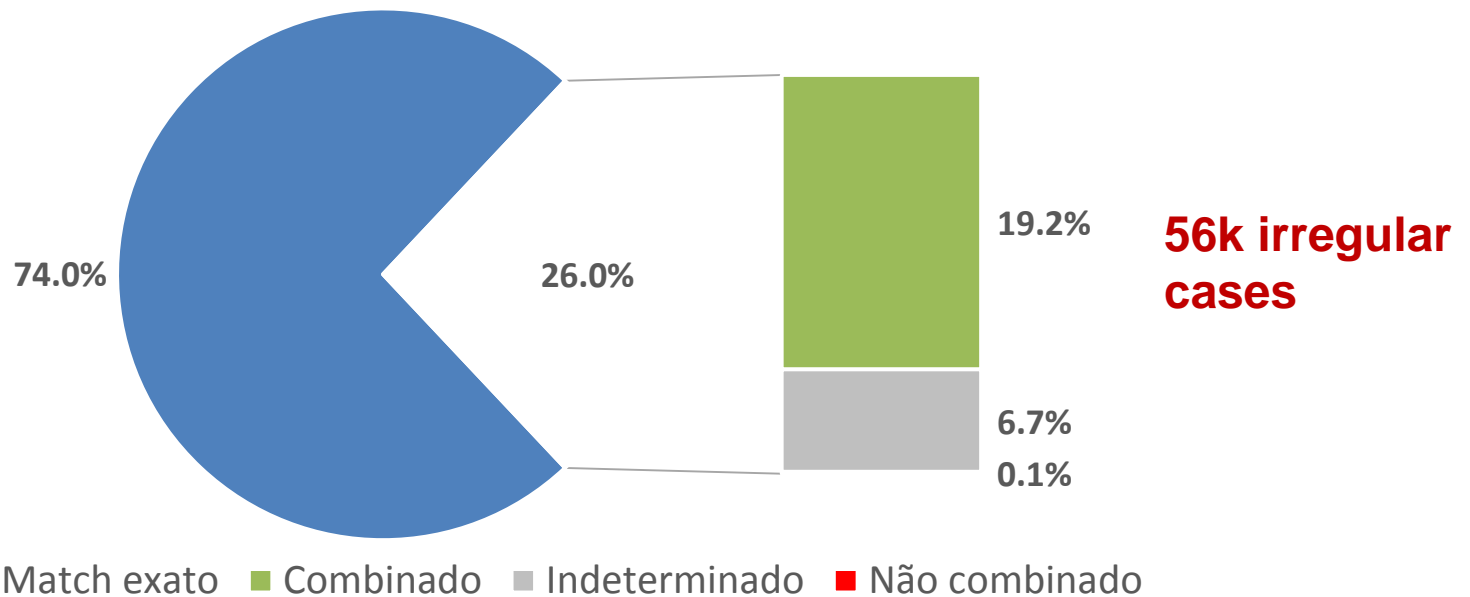
## Results – Labor

- Focus on Data Collection and ETL processes
- Implementation of 4 non-compliance typologies
  - Unemployment benefits for Artisanal Fishermen with incompatible income from other sources (private sector)
  - Unemployment benefits for Artisanal Fishermen with incompatible income from other sources (Pension benefits)
  - Unemployment benefits for Artisanal Fishermen with incompatible income from other sources (Brazilian Conditional Cash Transfer Program – *Bolsa Familia*)
  - Unemployment benefits for Artisanal Fishermen to deceased workers
- Possible savings of USD 1 mi (R\$ 4 mi)

# Cycle 2015


## Results – Social Assistance

- Implementation of 1 non-compliance typology
  - *Bolsa Familia* program benefits with incompatible income from other sources (data crossing)
- Possible yearly savings of USD 50 mi
- Data Quality test on the CadUnico (Social Assistance database)



# Cycle 2016

## Pensions

- Data Mining in Pension Fraud/Irregularities
  - Target universe
    - 4,359 frauded benefits
    - 21,789 irregular benefits
    - 242,927 regular benefits

33 mi total of population
  - Time frame
    - Aug – Dec
    - Current phase: Modeling
  - Preliminary results
    - “Flat table” with 3,203 indicators
    - 242,927+4,359 (fraud) / 242,927+21,789 (irregular) rows
    - Calculated by standard-deviations of subgroup (dimension x metric x unit of observation)

# Cycle 2016

## Pensions

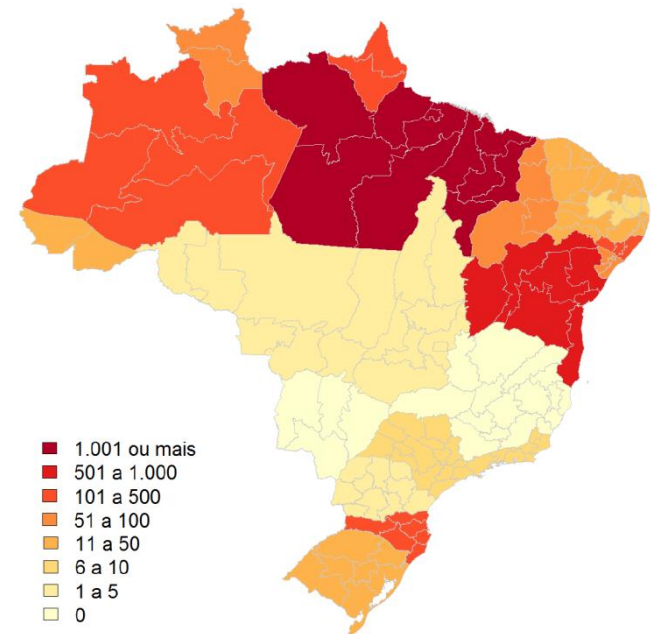
- Data Mining in Pension Fraud/Irregularities
  - Main challenges
    - Lack of information on old frauded benefits
      - Available data: jan/2014 – aug/2016
      - Original fraud cases of 29.538 reduced to 4.359
    - Poor Data Quality (high missing ratio on several columns)



# Cycle 2016

## Labor

- Automation of Data Collection, ETL and 4 previous typologies
- Implementation of illegal accumulation for Unemployment benefits (formal worker)
  - Artisanal Fishermen: 363 pymt (jan/16)
    - Concentrated in time (depending on the species) and region
  - Formal worker: 2 mi pymt (jan/16)
    - Up to 5 payments, nationwide





# Cycle 2016

## Social Assistance

- Data Quality/Enrichment of CadUnico
  - Electoral database (high quality, biometry)
  - Private sector employment information
- Implementation of typologies
  - Incompatibility between self-declaratory low income and wealth (societal participation, expensive vehicles)
  - *Bolsa Familia* program benefits with incompatible income from Government sources (pensions, unemployment, public contracts)

- **Contacts:**
- **Rodrigo Hildebrand**
- **NCAD – Center of Data Analysis and Information Technology**
- **Department of External Control – Social Security, Labor and Welfare**
  - **[rodrigooc@tcu.gov.br](mailto:rodrigooc@tcu.gov.br) – +55 (61) 3316-7971**
  - **[secexprevi@tcu.gov.br](mailto:secexprevi@tcu.gov.br) – +55 (61) 3316-7365**

**Thank you!!**