

MULTI-DIMENSIONAL ANALYSIS FRAMEWORK FOR DATA ANALYTICS DEVELOPMENTS

Picard Conference, December 09-10, 2021

Dr. Mathieu Labare, Data Scientist – Risk Management – Belgian Customs

Mathieu.labare@minfin.fed.be

Jonathan Migeotte, Coordinator – Risk Management – Belgian Customs

Jonathan.migeotte@minfin.fed.be

01100010110100101100101001
PROFILE

A project funded by the European Union's Horizon
2020 research and innovation programme under the
Grant Agreement No 786748.



Federal
Public Service
FINANCE

CUSTOMS AND EXCISES



Belgian Customs

Data Analytics & AI



Data analytics @ Belgian Customs

- Data Science for 10 yrs
 - SEDA 2.0 risk engine
- Data science inside Risk Management
→ Trained customs officers
→ Complementarity with risk analysis

Analysis Team

Querying

- ✓ Statistic & reports
- ✓ Generic analysis
- ✓ Support to risk analysts

Datamining Team

Advanced analytics

- ✓ Predictive modelling
- ✓ Multi-dim. Analytics
- ✓ Thematic oriented

NEW AI Team

Disruptive technologies

- ✓ A.I.
- ✓ Image processing
- ✓ Highly specialized

PROFILE research project

- Research project with 5 Customs Administrations
- Machine learning & data analytics for customs risk management

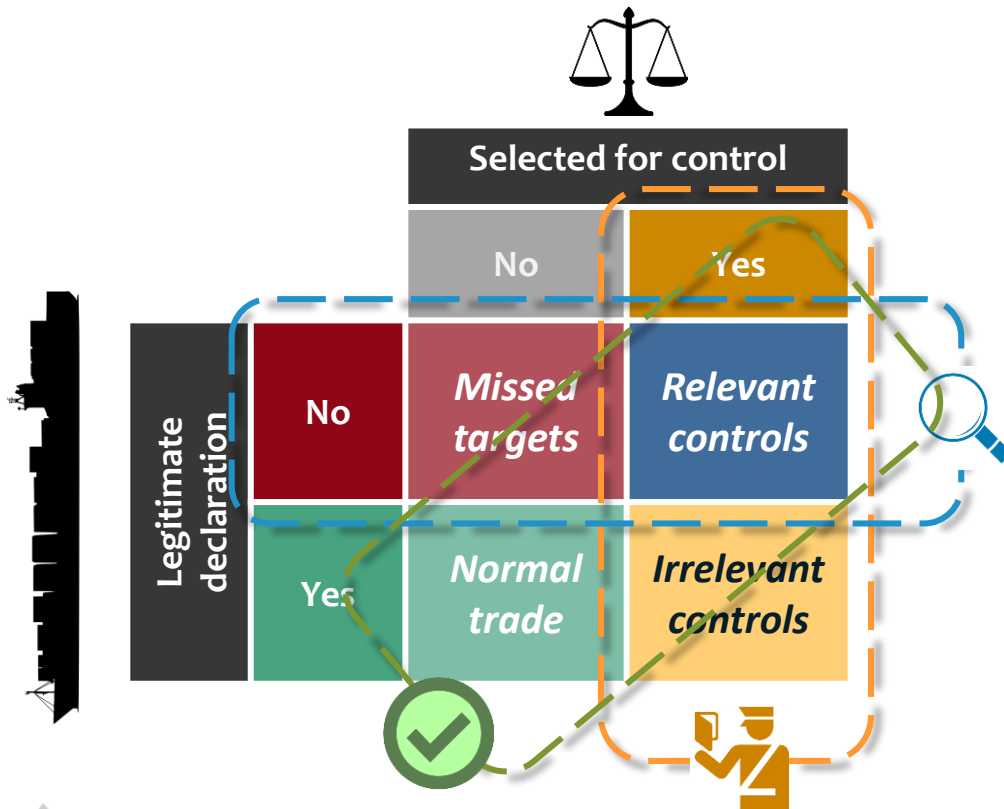
Belgian Living Lab

- Generic fraud detection
- Anomaly detection
- Summary declaration data enhancement



How to compare effectiveness of very different data analytics use cases ?

Efficiency through classical KPI's and their limitations



- **Precision** : *Percentage of good selections*

Problem : Does not quantify what is missed.

- **Sensitivity** : *Percentage of frauds found*

Problem : Estimation about what is missed must be available

- **Accuracy** : *Percentage of good decisions*

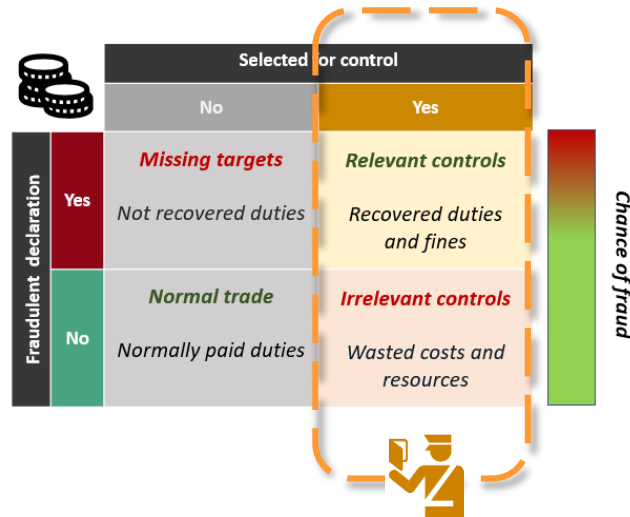
Problem : Customs data is strongly imbalanced
❖ *Accuracy will always be close to 1*

The myth of a global measurement of efficiency



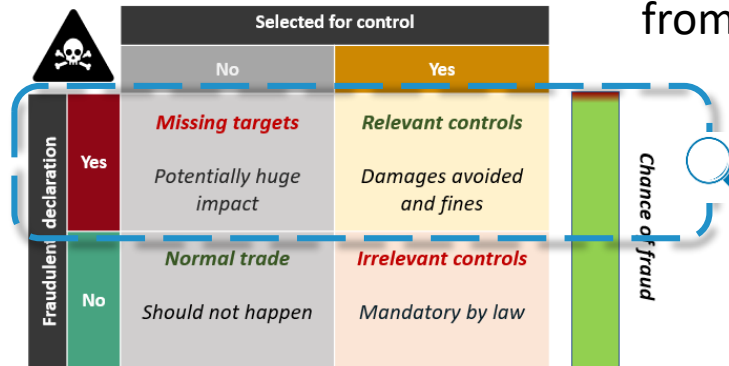
Financial Risks

- Optimize duties and taxes recovery
- Largest amounts have the most impact
- Finding each fraud is a waste of resources and money



Safety and Security Risks

- Prevent hazardous material from reaching the public
- Each risk has an equivalent impact
- Missing one fraud is potentially dangerous



Precision is *often* more important than sensitivity

Sensitivity is *often* more important than precision

Precision

Sensitivity



Analysis Framework for Risk Management Strategy



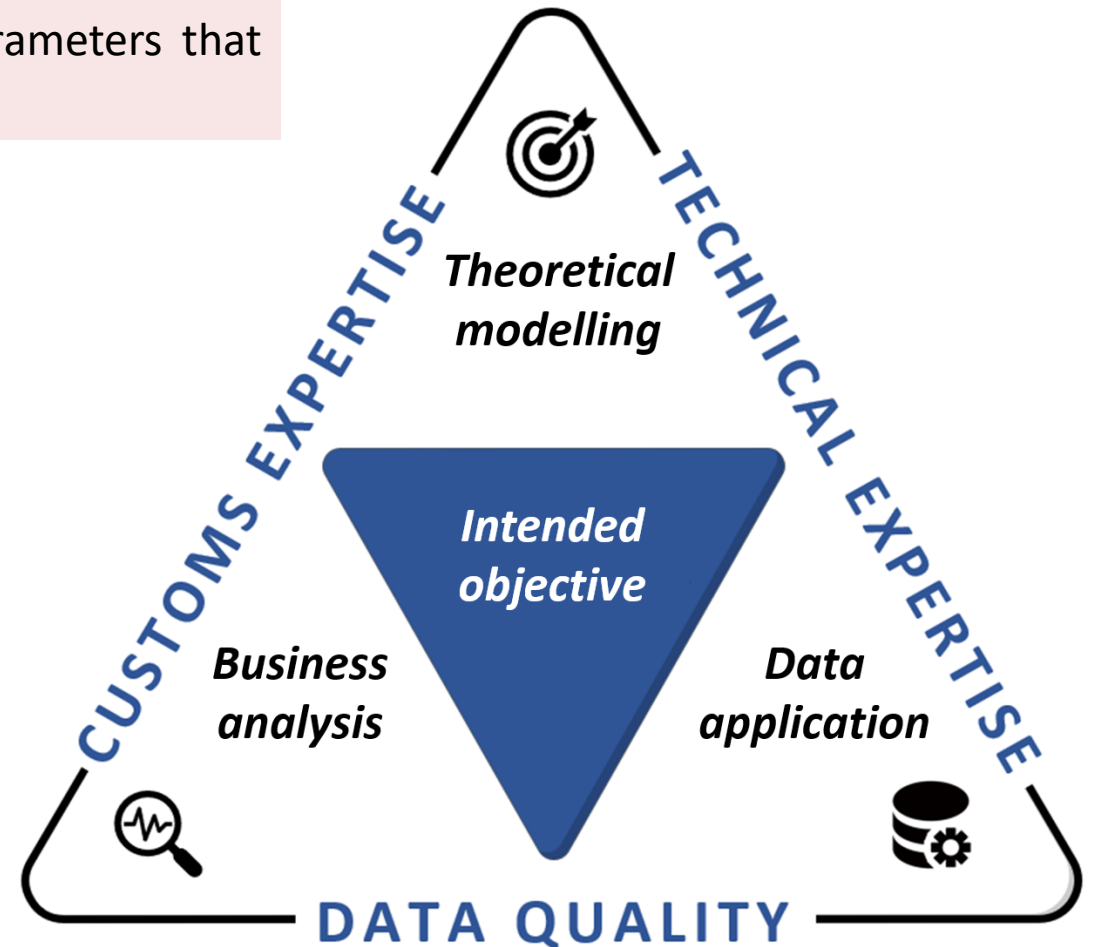
Generic framework of analysis that focuses on qualitative parameters that can be observed **in customs data analytics use cases**.

- apply to all sorts of results, **independently** of their level of development or readiness.

➔ What we lose in quantitative interpretation, is gained in the ability to effectively compare different results.

- provides an accurate analysis of the impact of the different investigated approaches.

➔ Highlight **specific strengths and weaknesses** and draw **lessons learned and recommendations**.



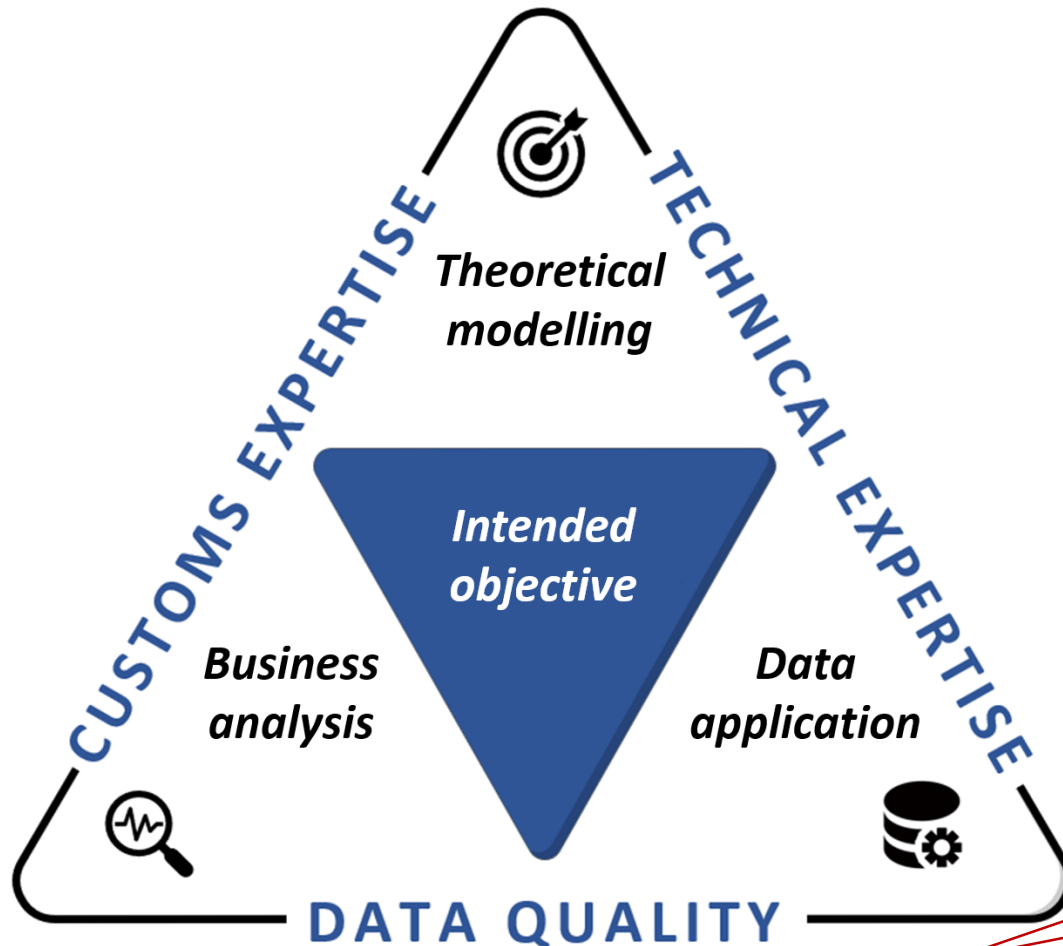
Analysis Framework

1st dimension : Data quality



Federal
Public Service
FINANCE

CUSTOMS AND EXCISES



Data is the foundation on which all the expertise can be developed and refined.

It must be:

- reliable
- structured
- up-to-date
- integrated

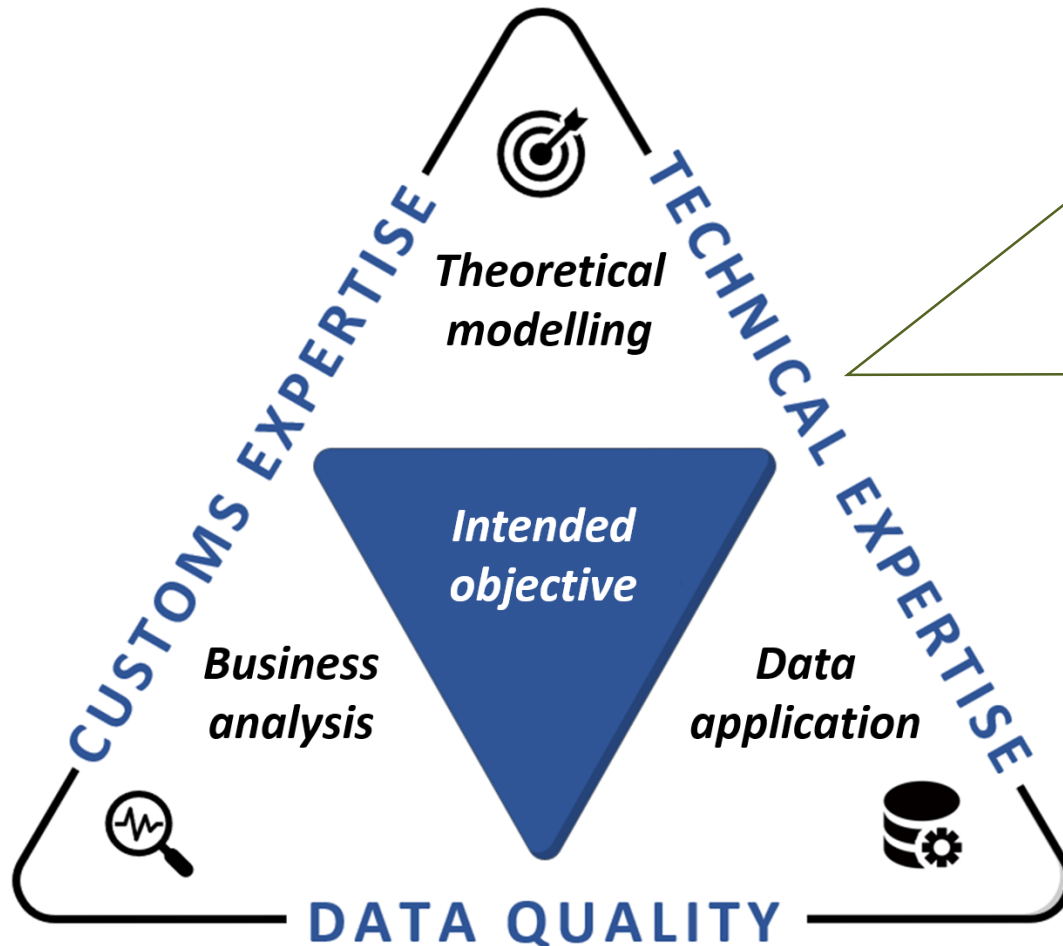
It should be handled with care:

- Confidentiality
- Data Protection
- 'AI Act'

Low data quality impacts every step of the process

Analysis Framework

2nd dimension : Technical expertise



- **Customs shifts towards digitalization :**

- Declarations and documents under electronic form
- Business procedure automatization
- Rapidly Increasing flows
- E-commerce

- **New skills and resources are needed :**

- IT infrastructures for data acquisition, storage and management
- Data analytics / Machine Learning / AI techniques for data handling and analysis
- Customs procedure automatization
- Methods and techniques evolve quickly

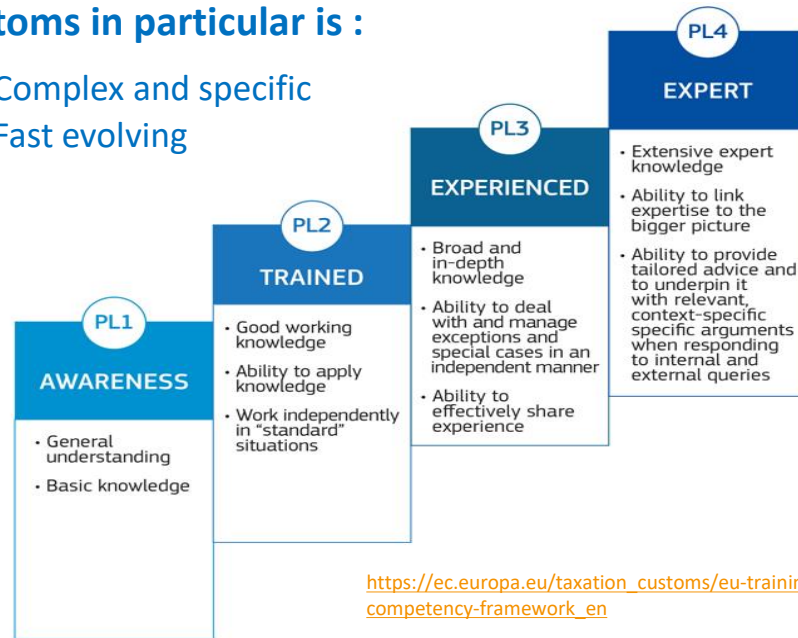
Customs must keep up with disruptive technologies

Analysis Framework

3rd dimension : Customs expertise

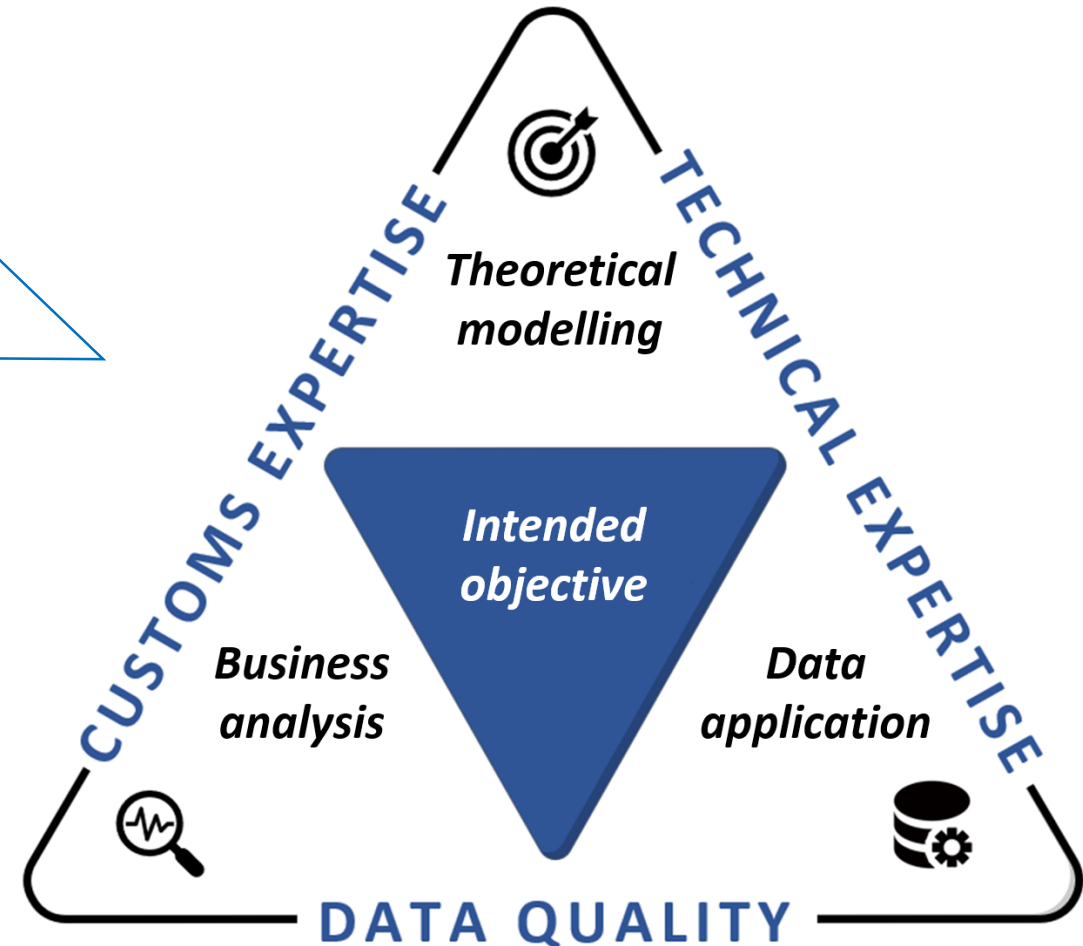


- The importance of domain expertise is underestimated
- Customs in particular is :
 - Complex and specific
 - Fast evolving



https://ec.europa.eu/taxation_customs/eu-training/eu-customs-competency-framework_en

Exploiting expertise is key in developing useful models



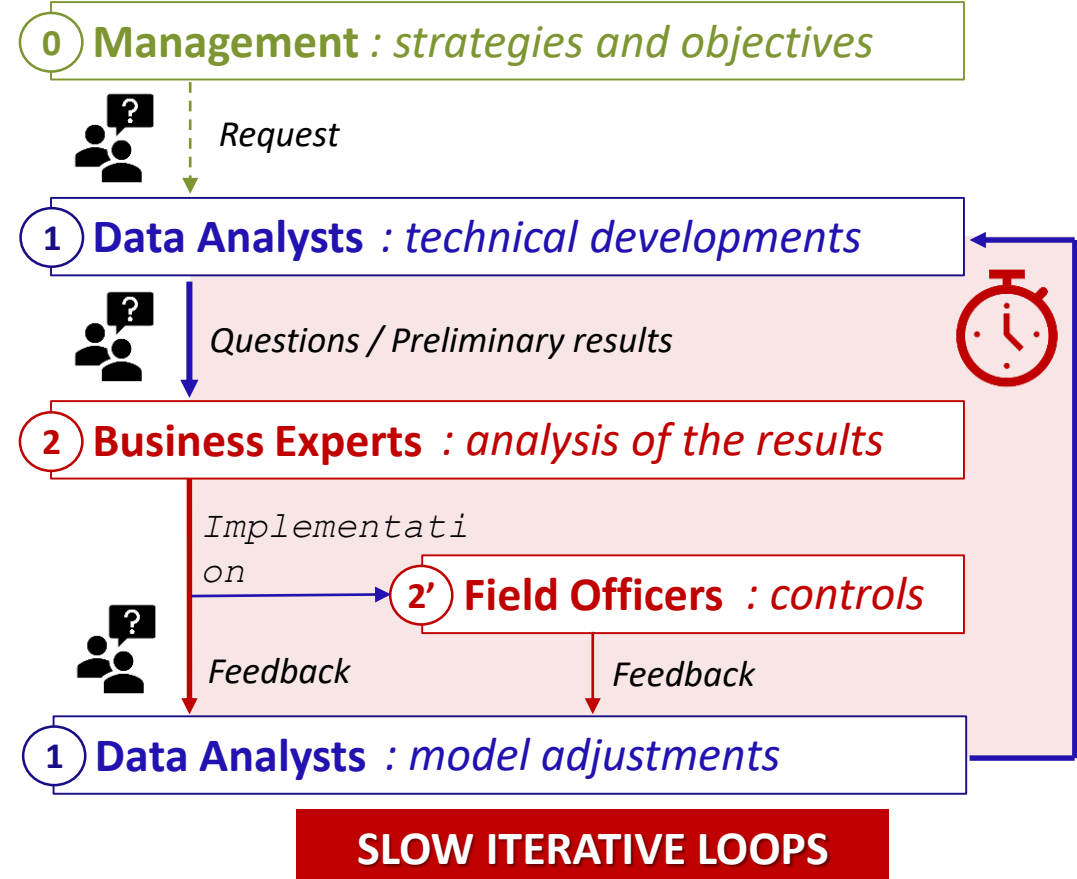
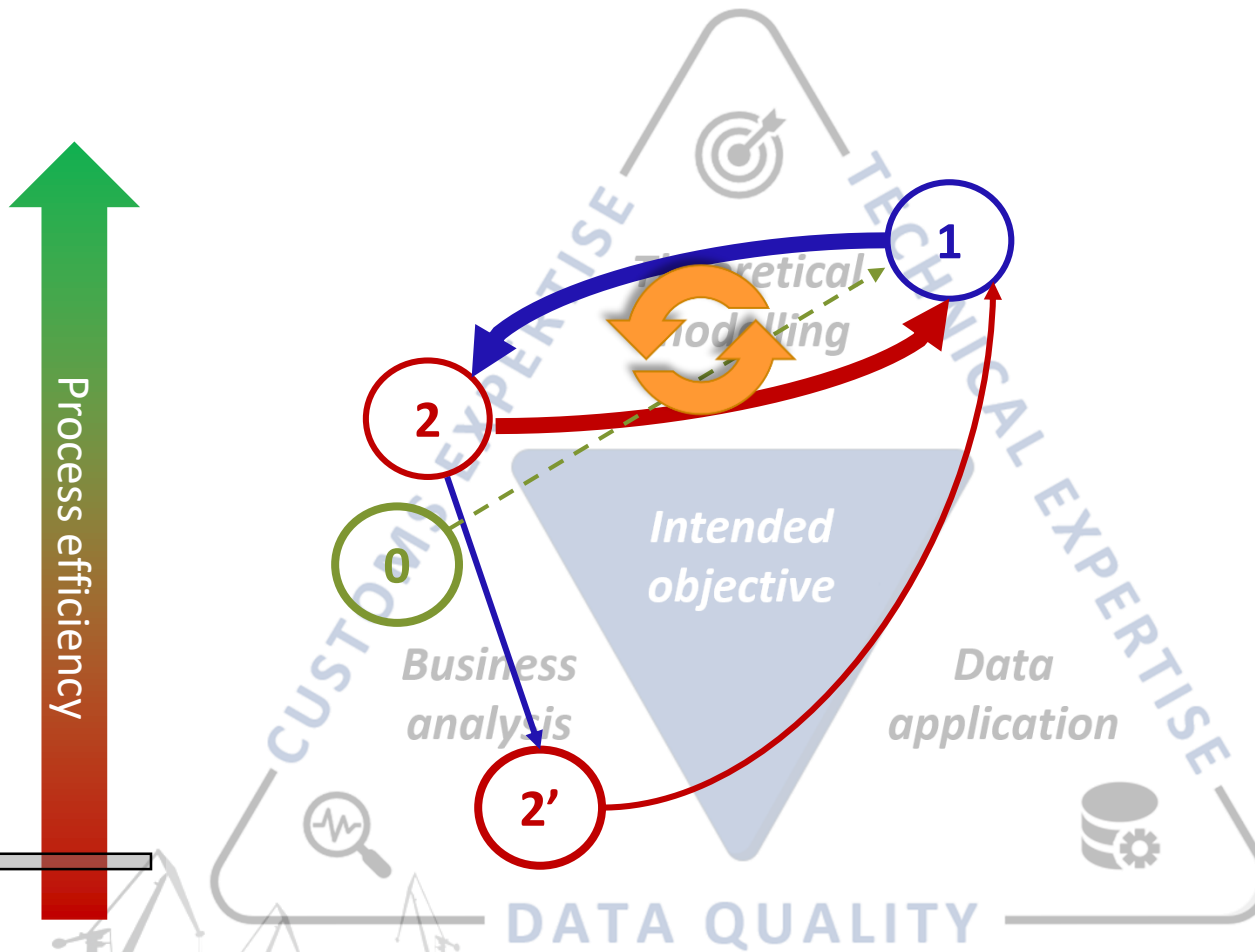
Analysis Framework

Working example : Implementation of a Data Analytics Process



Federal
Public Service
FINANCE

CUSTOMS AND EXCISES



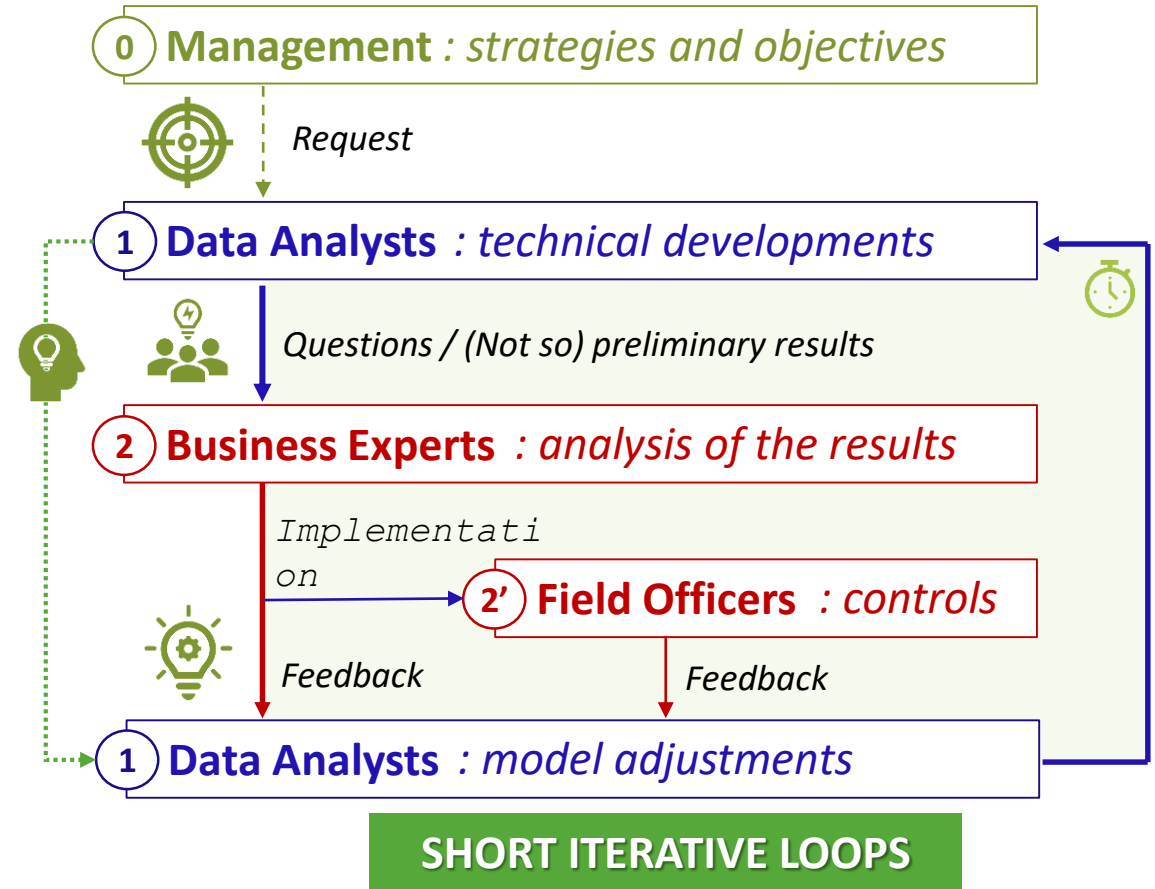
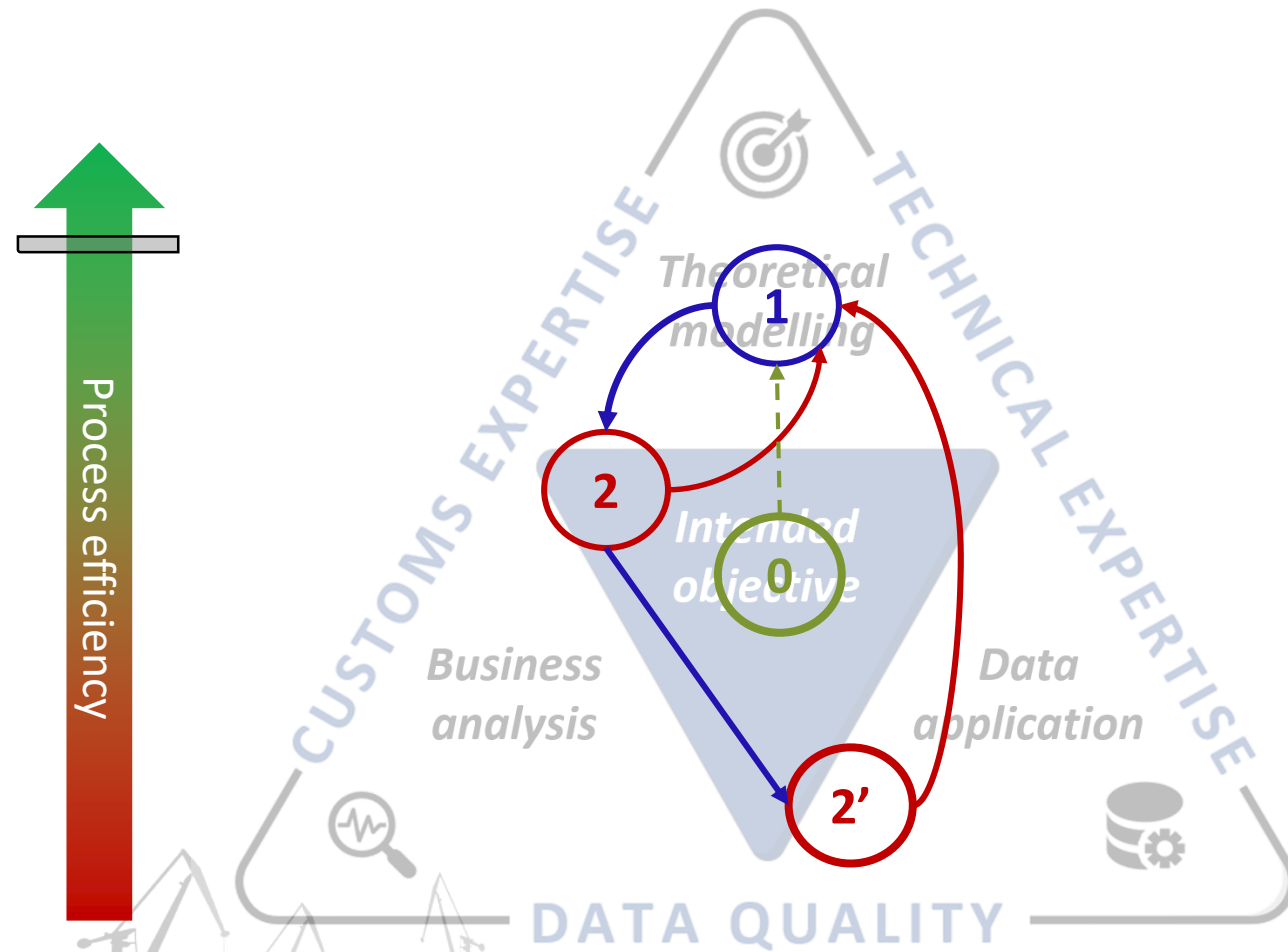
Analysis Framework

Working example : Implementation of a Data Analytics Process



Federal
Public Service
FINANCE

CUSTOMS AND EXCISES



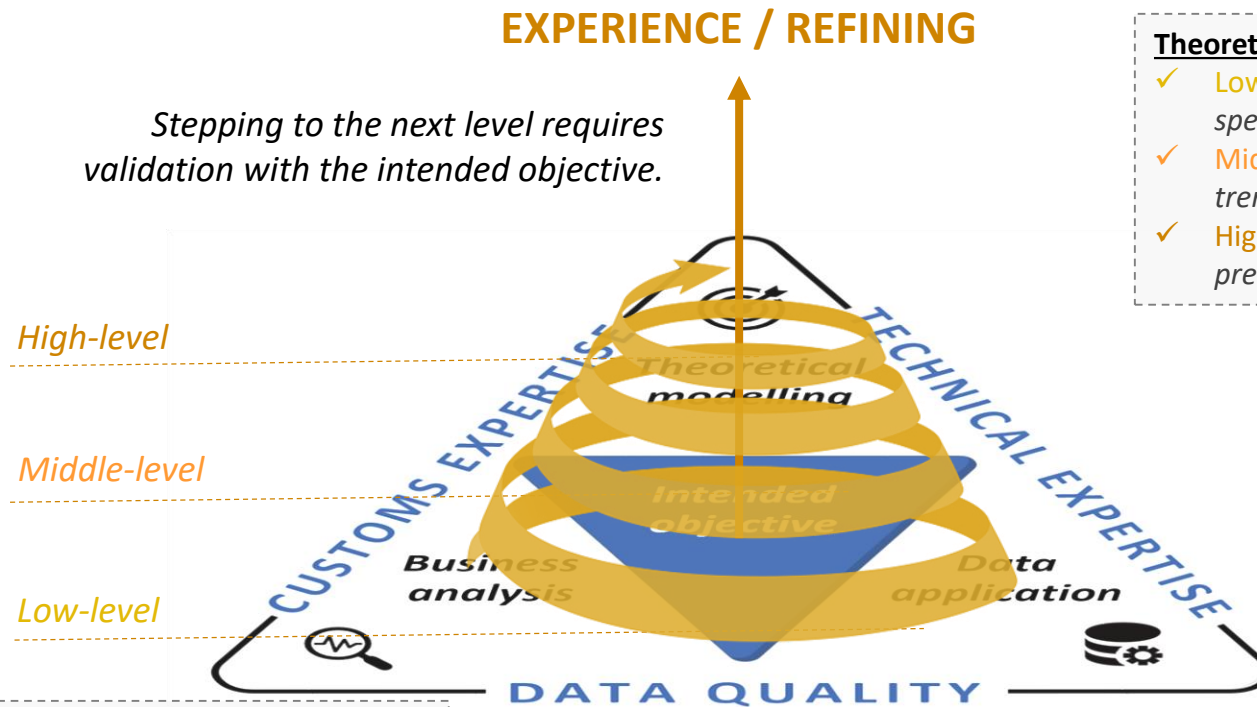
Analysis Framework

Towards a 4th Dimension : Iterative Learning and validation



Federal
Public Service
FINANCE

CUSTOMS AND EXCISES



Theoretical Modelling

- ✓ **Low level** : pre-processing based on global customs specificities
- ✓ **Middle level** : Identification of known or unknown trends based on exploratory analysis
- ✓ **High level** : Refined models in adequation with the pre-defined needs and objectives.

Data Application

- ✓ **Low level** : data cleaning / storage and checks for global inconsistencies
- ✓ **Middle level** : Enrichments and data-linking of homogeneous data
- ✓ **High level** : Complex data-linking, data scoring, dynamical implementation,...

Business Analysis

- ✓ **Low level** : Basic understanding of customs and legal framework
- ✓ **Middle level** : Advanced knowledge of domain specific patterns
- ✓ **High level** : Expert level knowledge

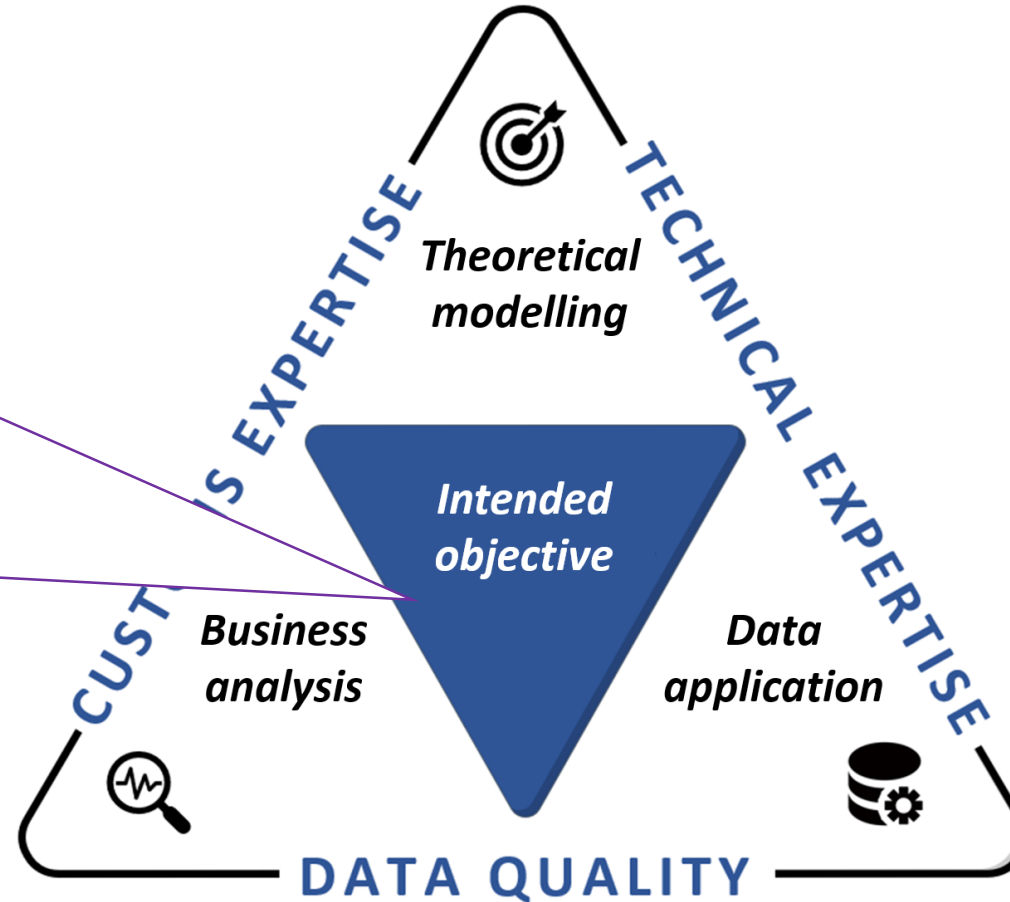
Ways Forward



Core dimension

Managerial Perspective :

- How to make sure that all dimensions are well articulated and are serving the objective ?
- What impacts the 'gravitational pull' of the objective and how to improve it?

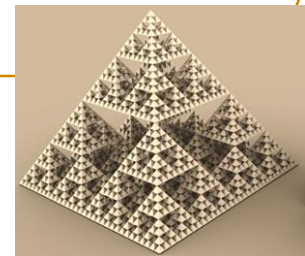


How generic is it ?

- Mainly applied to Customs Risk Management
- Other relevant contexts ?
 - Customs related
 - Other domains

Levels of abstraction

- Focus on practical experiments until now
- What if we look at it from further away ?



MULTI-DIMENSIONAL ANALYSIS FRAMEWORK FOR DATA ANALYTICS DEVELOPMENTS

Summary & Take-Home Messages

- Extension and validation in progress
 - Developed in a Customs context → to be applied in the Profile Project
 - Positive feedback from other Customs Departments
- Application to others domains to be confirmed (by practice)
- Expertise and knowledge exchange is key for optimal resource usage

01100010110100101100101001
PROFILE

A project funded by the European Union's Horizon
2020 research and innovation programme under the
Grant Agreement No 786748.



Federal
Public Service
FINANCE

CUSTOMS AND EXCISES

