

# Big Data and ML / AI @ Altice Labs

**Luis Cortesão**

Innovation and Technology Coordination

**Seminário Mestrado em Análise e Engenharia de Big Data  
DI-FCT-UNL**



# About Altice Labs

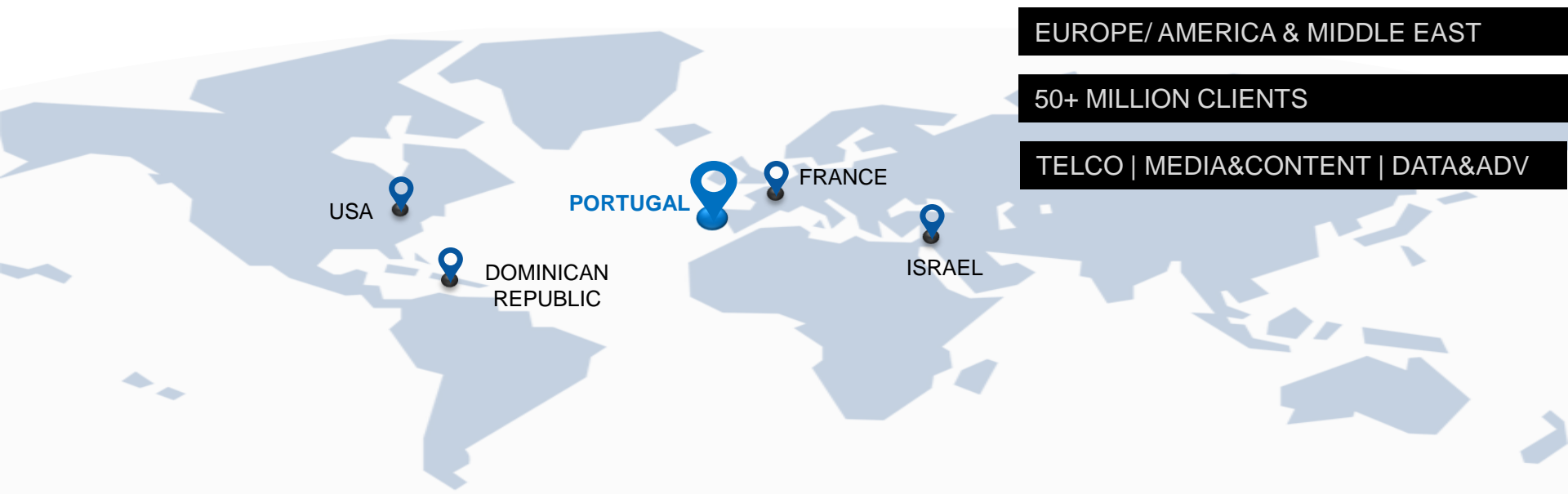
# Altice Group

## Altice Portugal

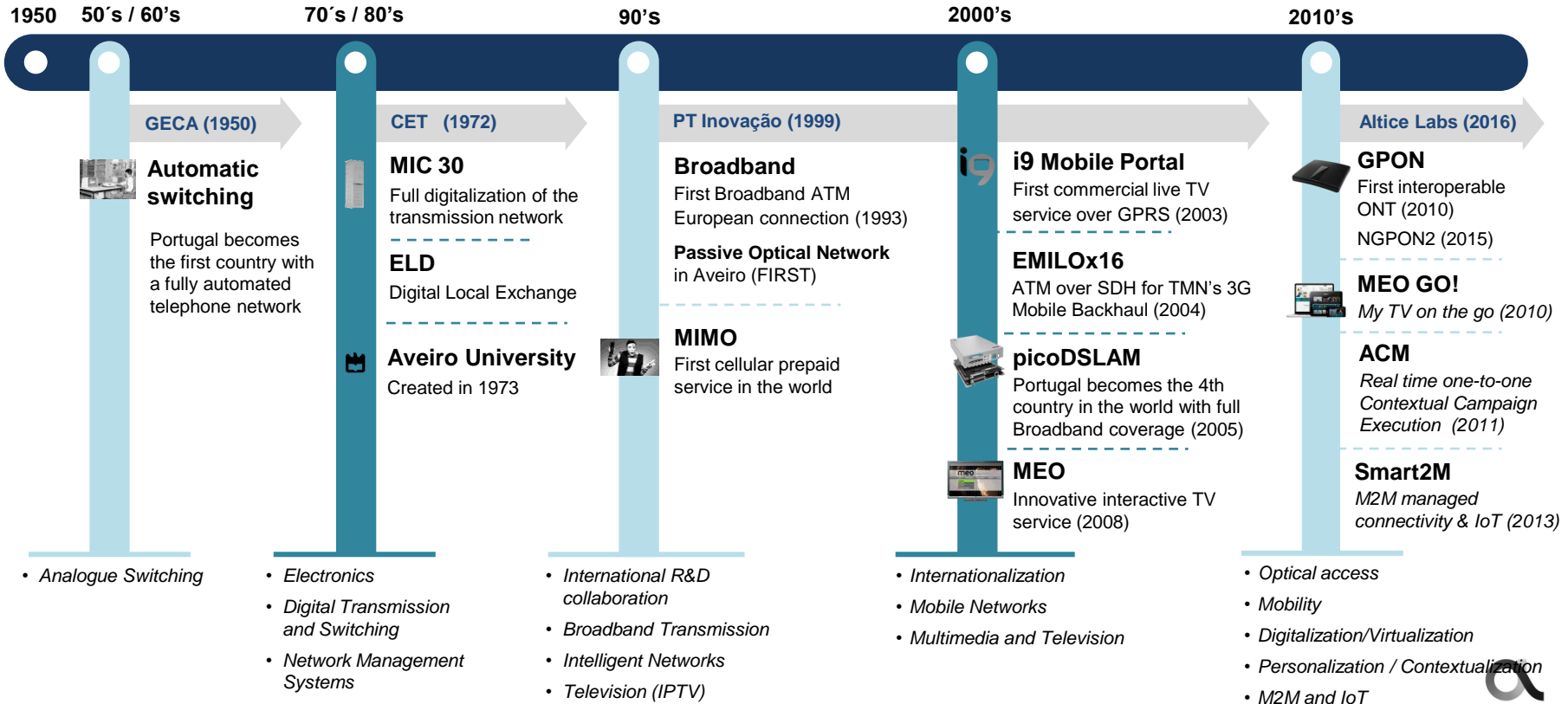
Portugal telecommunications market leader, with a vast service portfolio, ranging from fixed and mobile voice and data services to Cloud & IT, TV & media.

## Altice Labs

Recognized innovation background with several RDI centers



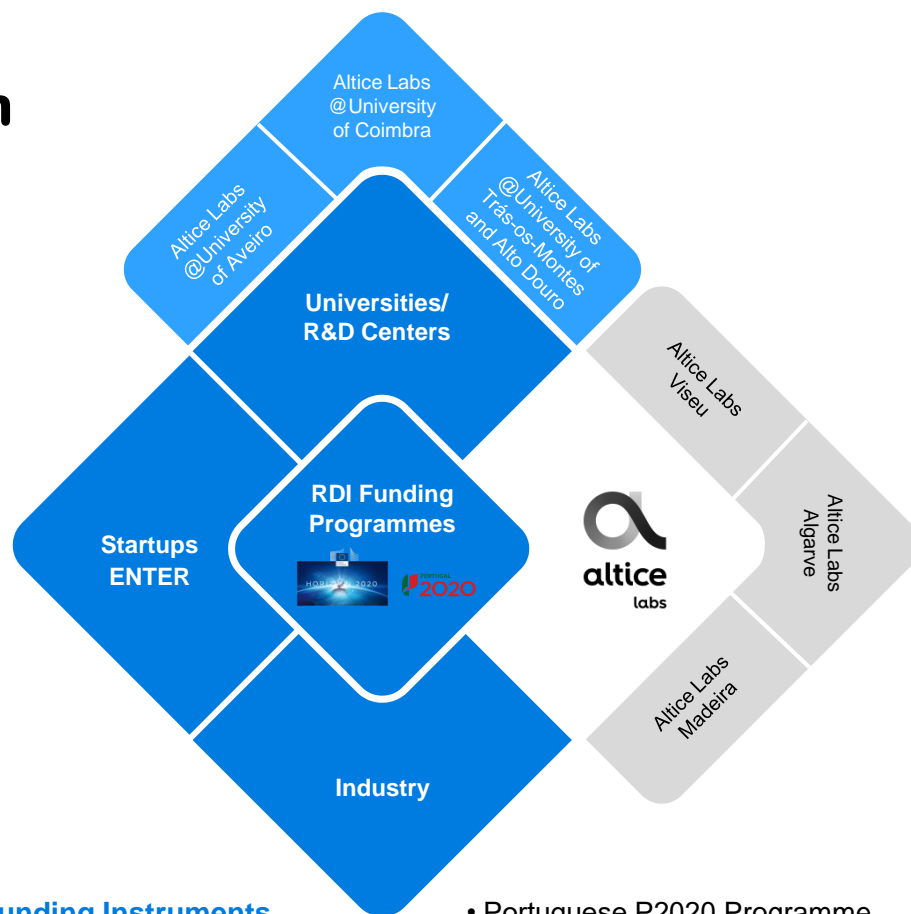
# The history of Altice Labs is linked to the Portuguese telecommunications sector evolution



# Innovation Process & Ecosystem



Altice Labs continuously engages in collaborative RDI projects as part of a sustained strategy for technological leadership.



## Funding Instruments

- European FP7 Programme
- European H2020 Programme

- Portuguese P2020 Programme
- Universities
- Startups, Manufacturers and SMEs

# Space ENTER

## Altice Labs - Lisboa

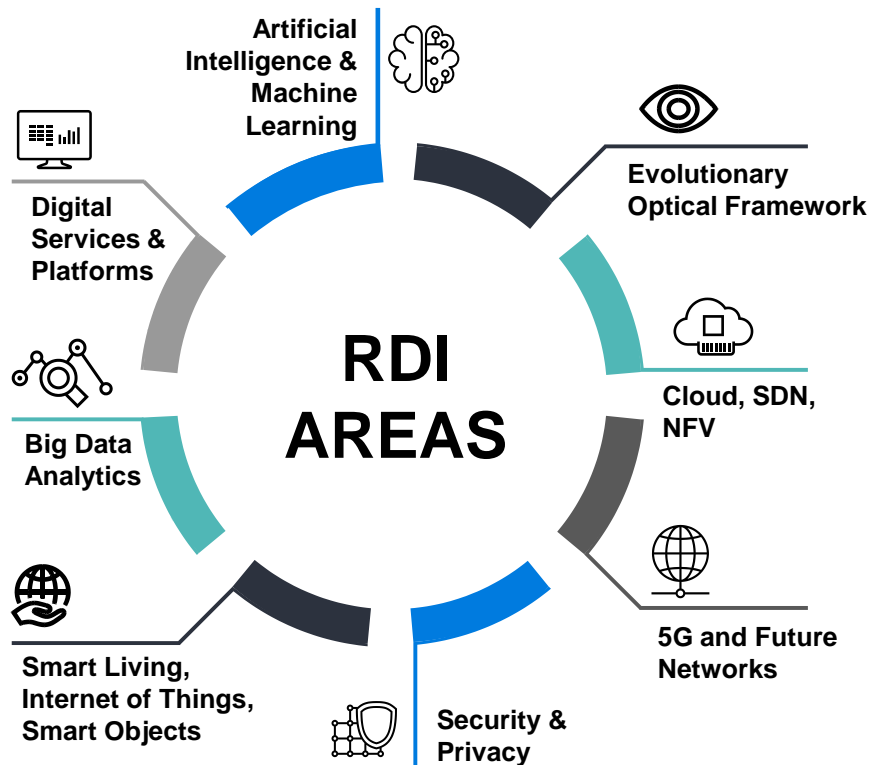
- 800 m2 in a trendy zone
- Near several services and transports (metro, train, boats and buses)
- Co-working space, snack, meeting rooms, auditorium, laboratories
- Equipped with Altice Labs and partners equipments, allowing excellent working environment
- Intended for startups working with Altice Labs but may be used by third parties (temporary and subject to conditions)
- It can be used by partners to host events (workshops, master classes, lectures, etc.)



# RDI Areas



To ensure value creation and positive return to the operations, Altice Labs guarantees that knowledge transfer occurs in all exploratory innovation specific activities around strategic areas.



# Bussiness Units



## Network systems



- Optical Network Equipment
- FTTx Networks
- G.fast
- Certification and Interoperability Labs
- Cellular Planning and Optimization Services



## Operations support systems



- Network Development and Inventory
- Service Fulfillment
- Service Assurance
- Customer Centric Operations



## Network control & services platforms



- Convergent services
- Charging and policy management
- Campaign and channel management
- M2M
- API Management



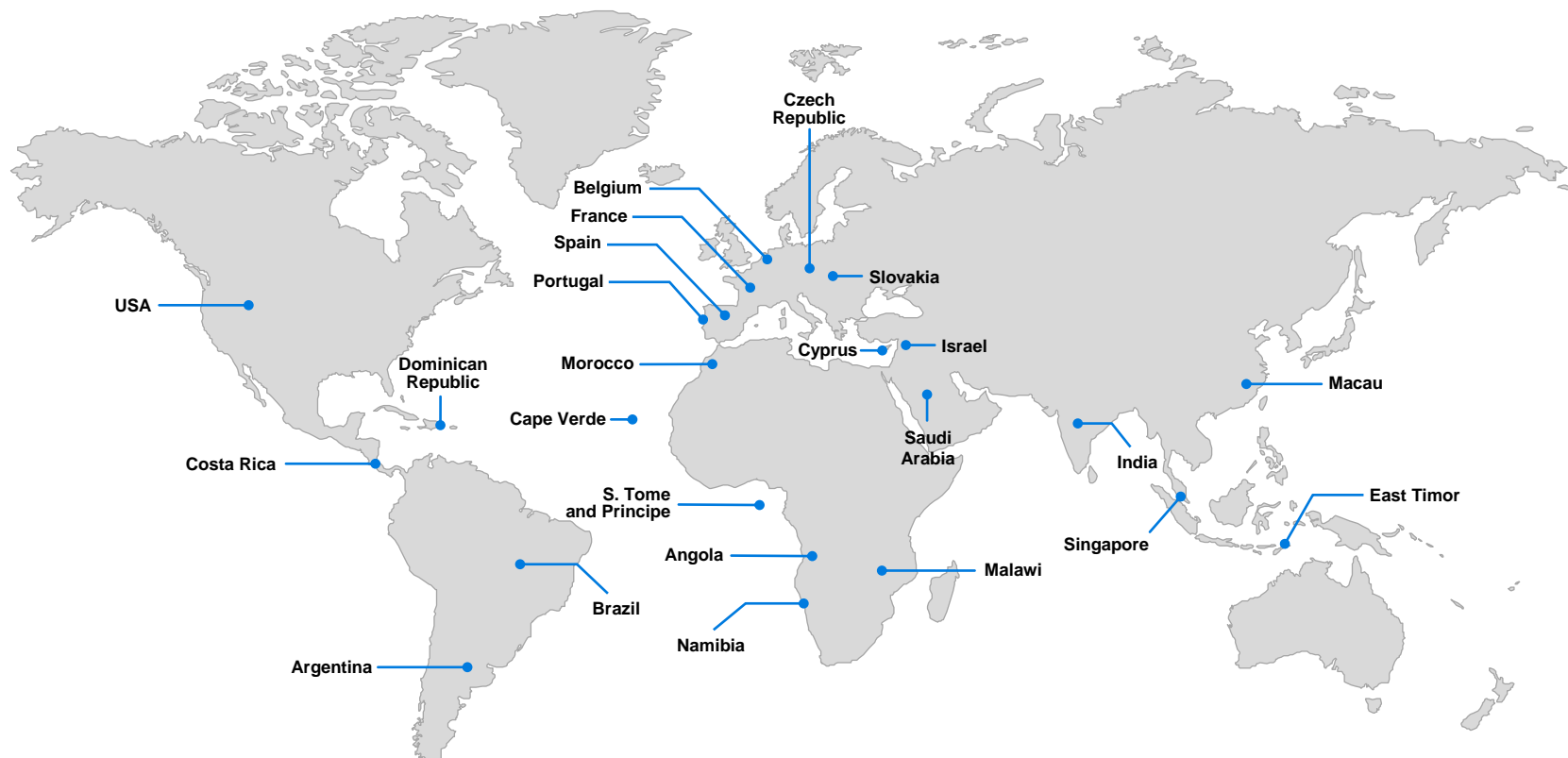
## Digital, internet & television



- Interactive TV and OTT
- Wallet & Payments
- Push&Pull advertising platforms
- Personal Cloud Storage
- eCare & eHealth

Professional engineering services, technological training and consulting

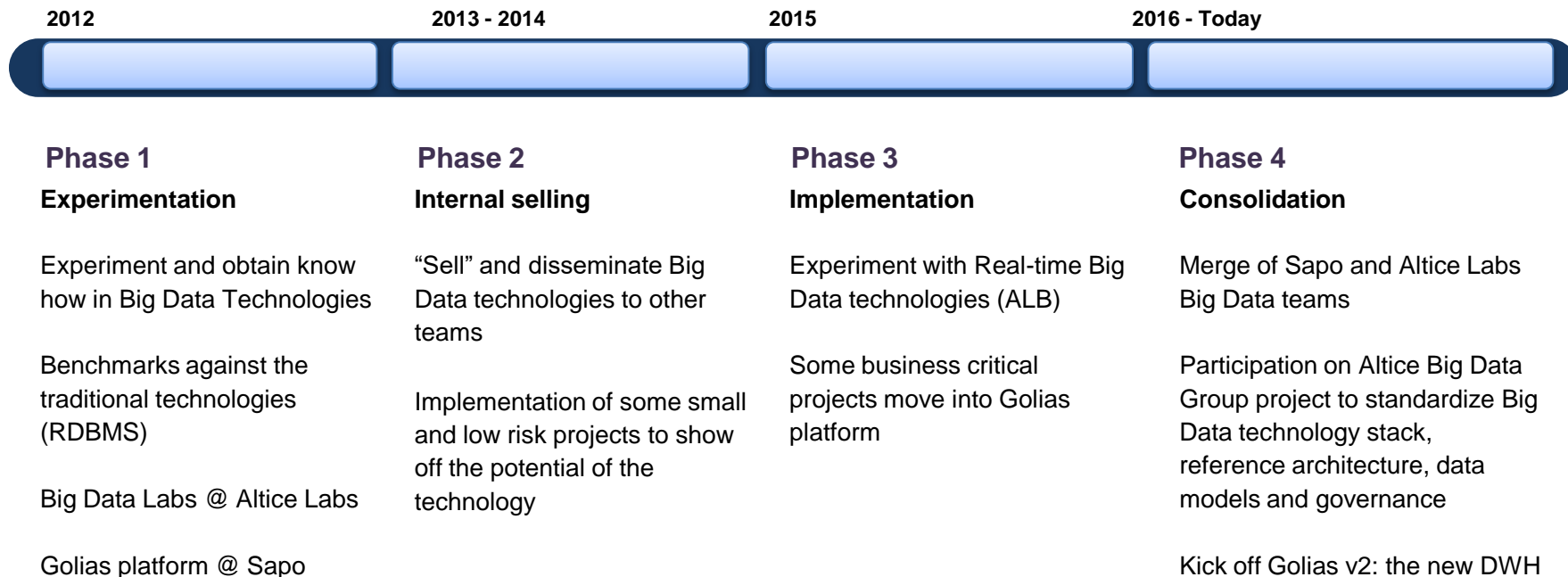




250 million people communicate everyday through technology developed by Altice Labs

# Big Data @ Altice Labs and MEO

# Big Data @ Altice Labs and MEO



# About the “Golias” platform

## NUMBERS

Total of 71 servers holding 281 TB compressed data (2.5 PB uncompressed) and 12 TB (uncompressed aggregated data in business specific Data Marts) divided by 108 databases in 10 different clusters

Growing at 10 TB (uncompressed) by month

Data is received from +40 different distinct systems and is processed by +240 distinct processes or services

## SOFTWARE USED

Cloudera for the Hadoop clusters, 3 PostgreSQL clusters, 2 Cassandra clusters, MonetDB, and 49 instances of Solr

Data is handled by Pig, Java and Python processes

## DATA INGESTION

By batch processes that import data from external databases

By Sapo Broker, and in house developed pub/sub messaging platform

By service listeners that receive events from webpages, mobile apps, set top boxes apps, etc

## DATA ACCESS

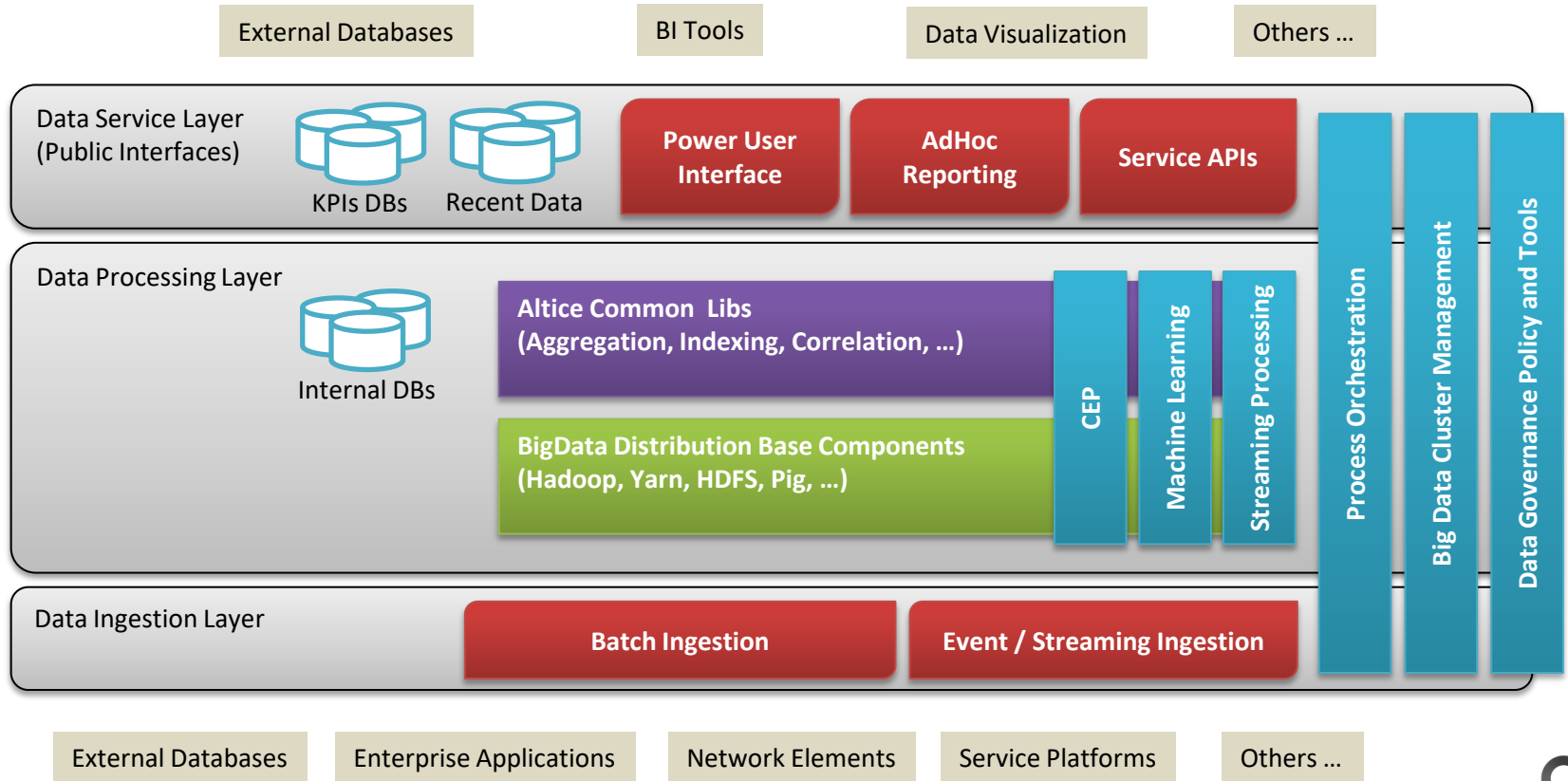
In house developed web applications for analytics and dashboards

Using in house developed API interfaces

Ad hoc reports as needed by the business areas

# Altice Big Data reference architecture

# Reference Architecture



# Technology Stack

External Databases

BI Tools

Data Visualization

Others ...

Data Service Layer  
(Public Interfaces)



Data Processing Layer



Flink



Data Ingestion Layer



External Databases

Enterprise Applications

Network Elements

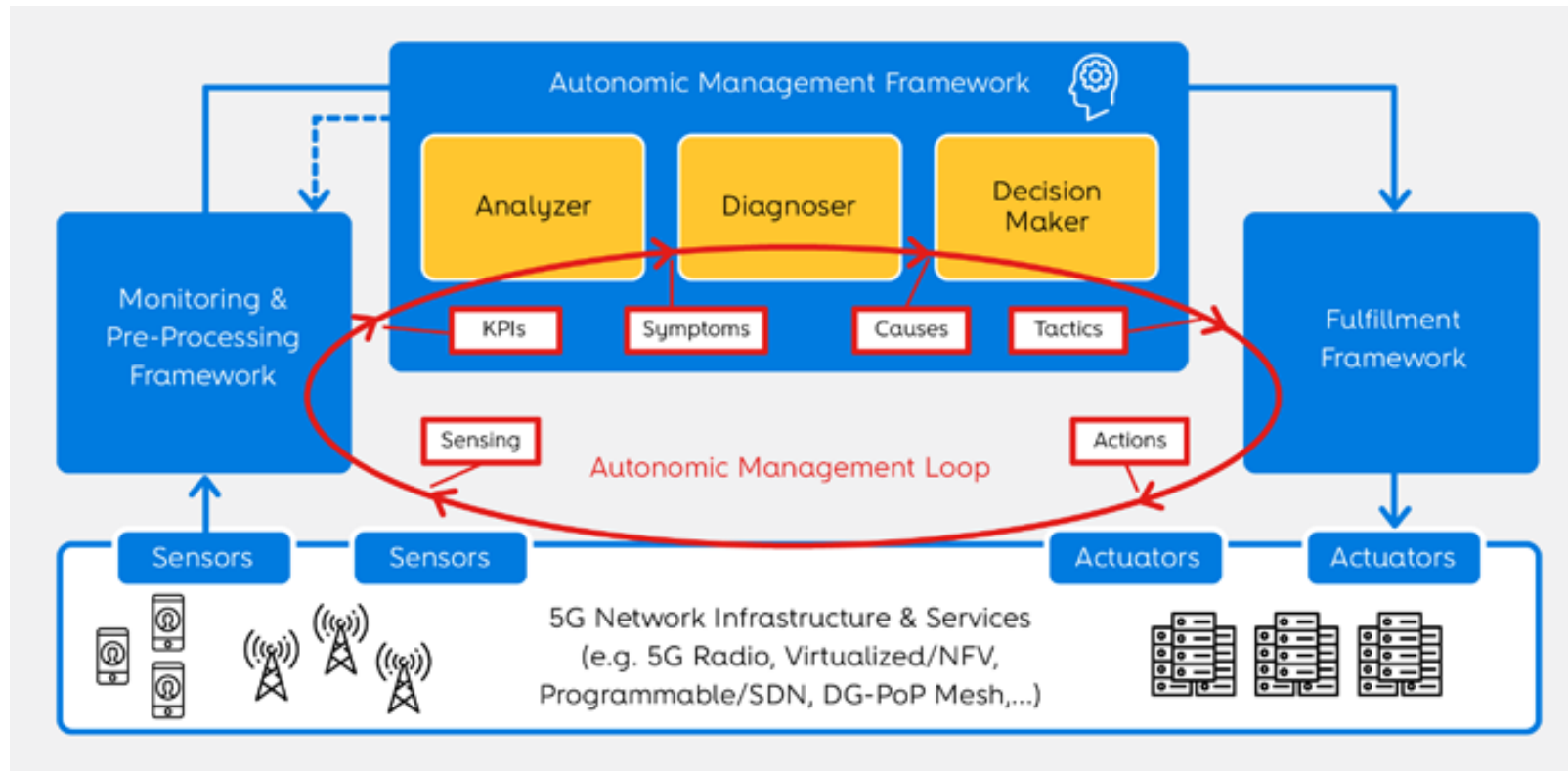
Service Platforms

Others ...

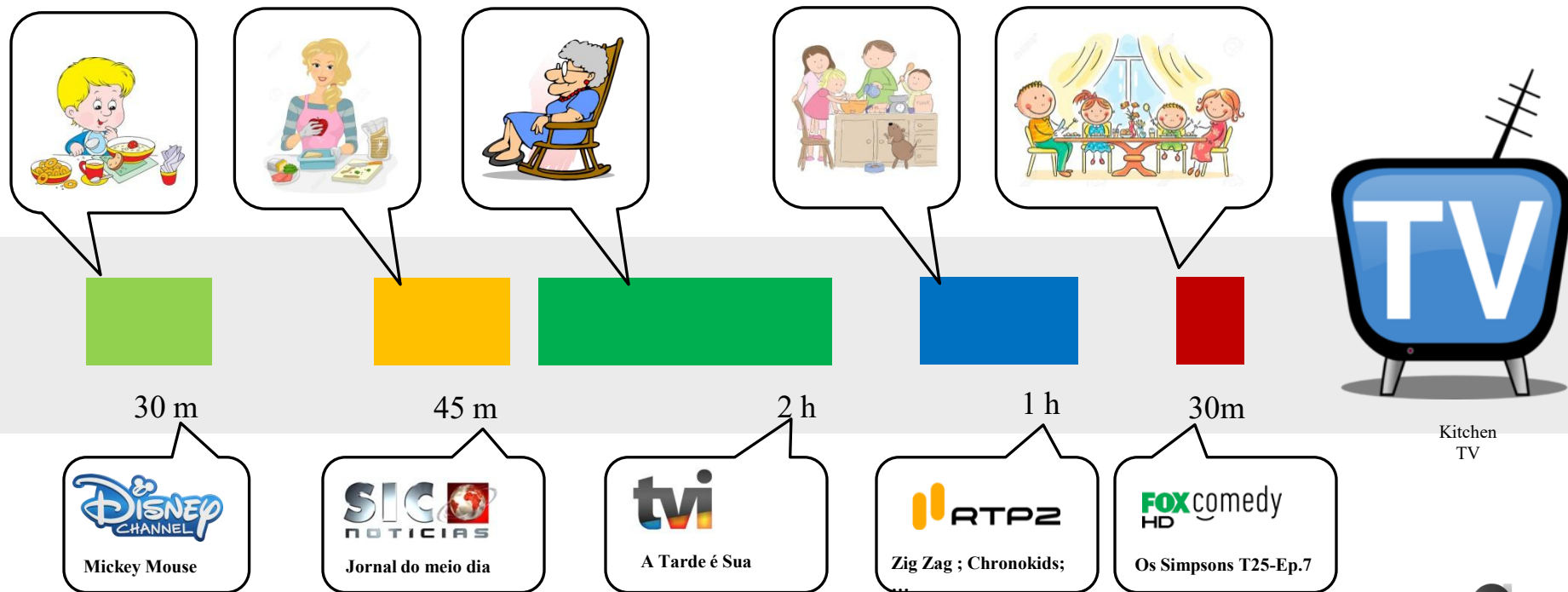
# Some projects using ML / AI



# H2020 (5G-PPP) SLICENET Project



# Content recommendation system



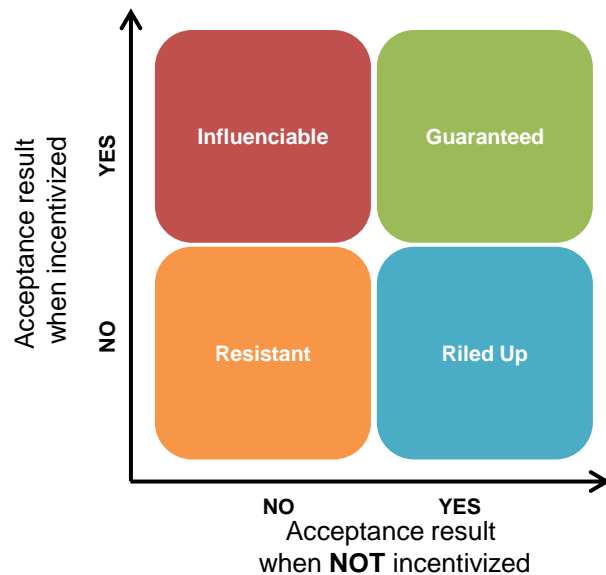
# Campaign Management

Campaigns are a common methodology used by Telecom Operators to improve revenue and customer satisfaction, but if done wrong it can backfire

This project aims to answer two questions:

1 – Avoid to annoy users that won't be receptive to a particular campaign (Resistant and Riled Up)

2 – Identify the users that telecom operator can influence with a particular campaign. I.e., change the normal behaviour of that user (Influenciable)



# Final Notes

# Final notes

## **Huge ML and AI momentum**

- Everybody wants to be part of it and try to use ML ... (even when it's not the best option or the technology is not yet good enough)

## **Telecom operators generate very large amounts of data**

- And everything is much more difficult (it's not for your average python script)

## **Human resources with good knowledge of ML still hard to find**

- Training them takes much time and you have no guarantees they will stay around for long time

## **The progress so far**

- None of these projects reached the end yet, but they have been useful as a way to learn and experiment

# Genius Program

Altice Labs provides in this context, Innovation projects that rely on your talent and capabilities.

<http://www.alticelabs.com/en/challenges.html>



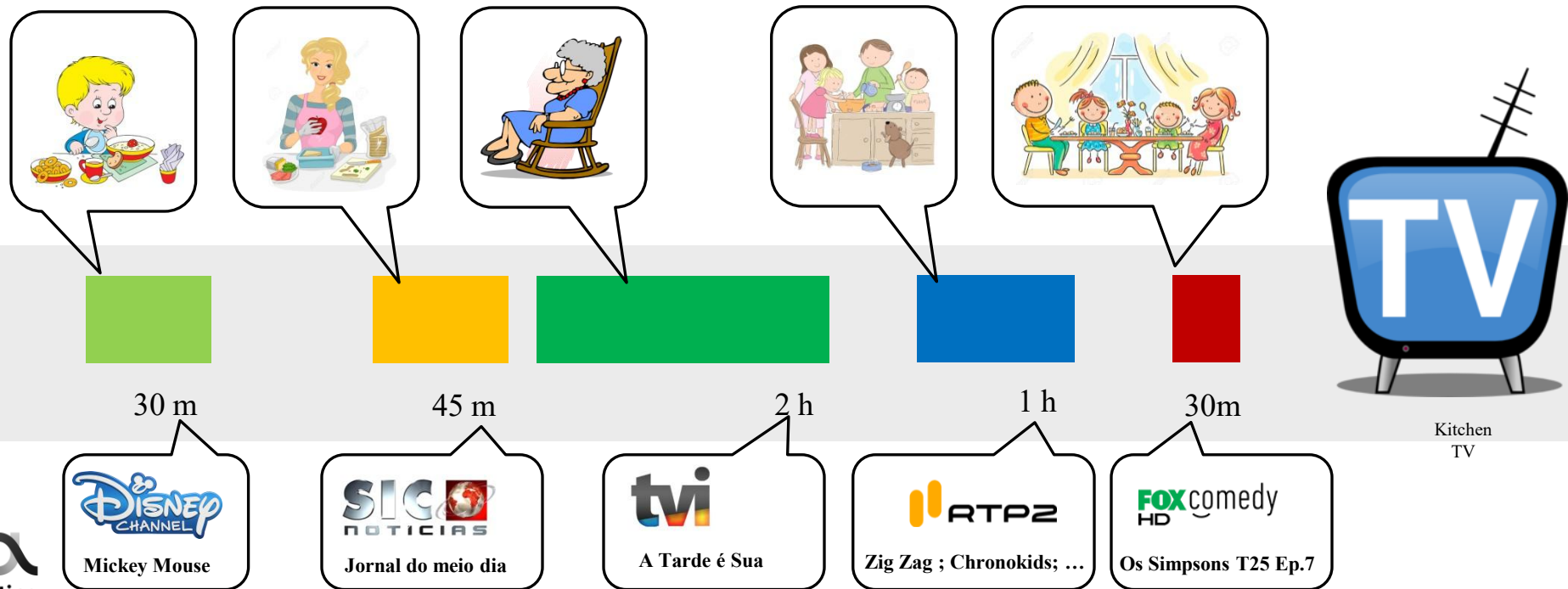
# MEO IPTV Recommendation Scenario

# INDEX

1. O contexto
2. O problema
3. Dados disponíveis
4. Voltando ao problema

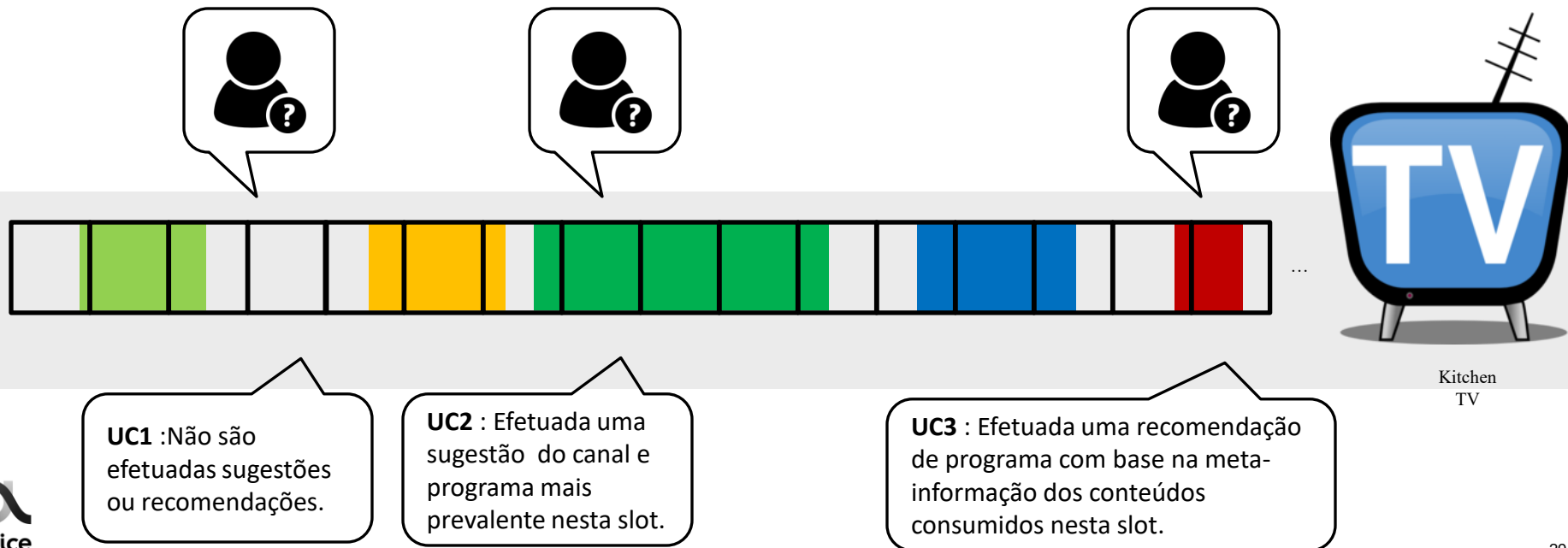


# 1. O Contexto



# 1. O Problema

Para uma determinada **STB, dia e hora**, ser capaz de realizar **uma sugestão/ recomendação** para o canal e conteúdo que terá mais alta probabilidade de ser escolhido.



## 2. Dados Disponíveis

- Os dados referem-se às duas seguintes atividades reais:
  - **Visualizations**

Esta atividade ocorre sempre no âmbito de uma determinada STB ( Set Top Box) quando um utilizador acede a um programa e mantém a visualização desse programa por mais do que **30 s.**
  - **Exhibitions**

Esta atividade ocorre sempre no âmbito de um canal e corresponde à disponibilização em stream de um determinado programa
  - **Programa**

Um programa de televisão descreve um conteúdo televisivo único, que pode ser exibido múltiplas vezes no mesmo canal e em múltiplos canais.
  - **Canal ( channel\_mappings )**

Stream de conteúdos televisivos, programas que são exibidos em sequência,

## 2. Dados disponíveis

### Tipo de conteúdo (genres)



## 2. Dados Disponíveis

### STB

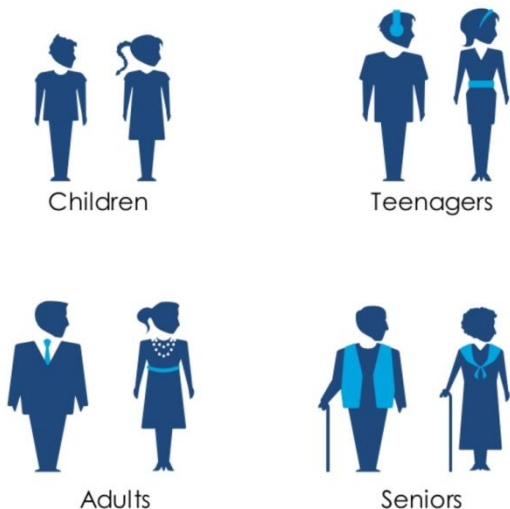
Dispositivo set top box que contem a lógica de menus de interface com o utilizador e recebe informação do comando remoto, do seu próprio interface de comandos locais e da rede IPTV para realizar as visualizações de programas e seus canais.

O identificador de cada box faz parte dos dados das visualizações.

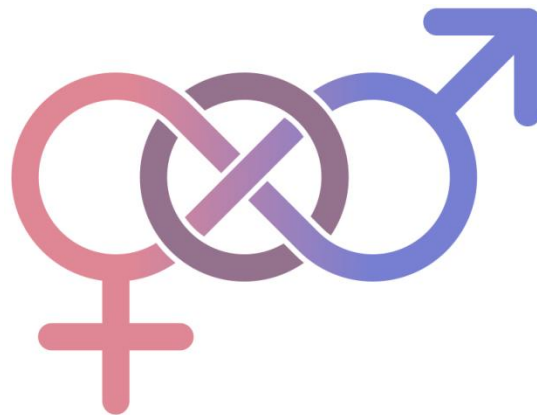


## 2. Dados disponíveis

Possíveis enriquecimentos a realizar (via canal) sobre os dados disponíveis.



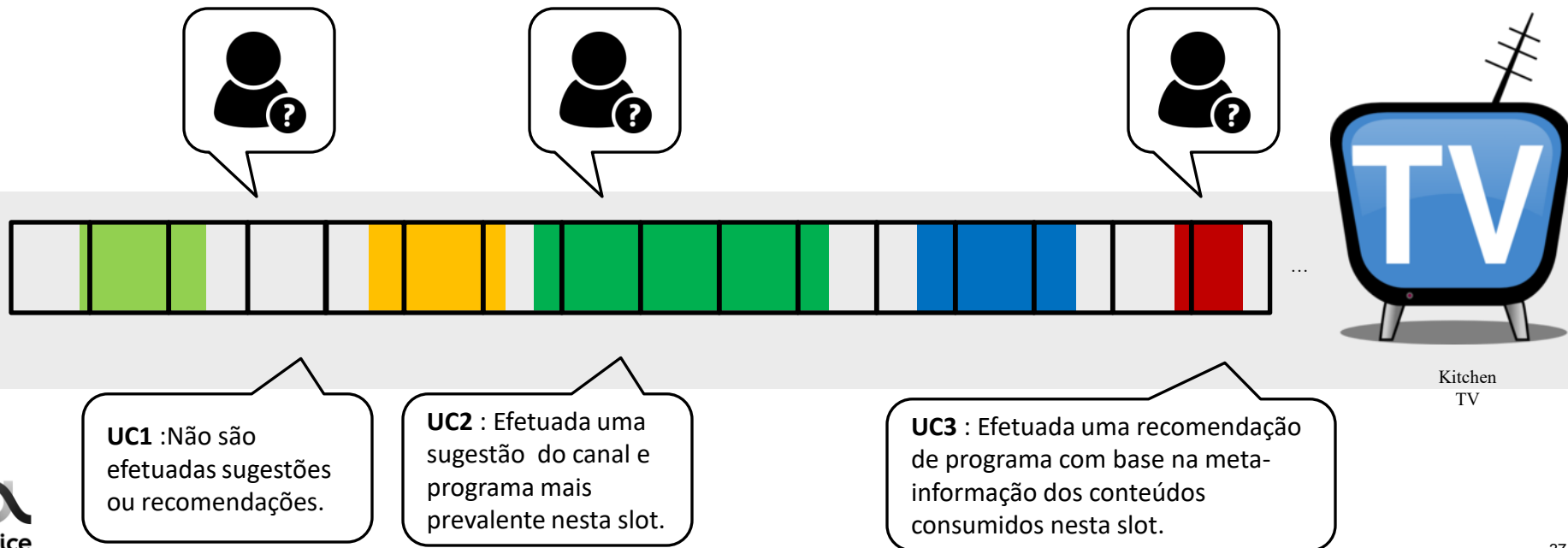
Segmento etário do público destino



Género típico do publico de destino

# 1. O Problema

Para uma determinada **STB, dia e hora**, ser capaz de realizar **uma sugestão/ recomendação** para o canal e conteúdo que terá mais alta probabilidade de ser escolhido.



# Thank You

## Any questions ?

Rua Eng. José Ferreira Pinto Basto,  
3810 - 106 Aveiro Portugal  
T: +351 234 403 200  
F: +351 234 424 723  
[www.alticelabs.com](http://www.alticelabs.com)