



BUPI BALCÃO ÚNICO
DO PRÉDIO

Big Data Analytics Expertise Centre

BUPI
Laboratório de Análise de Dados

MAR.2018

alTran

Agenda

- 1 Altran – o que fazemos?
- 2 O que é o BUPi?
- 3 Dados & Desafios iniciais
- 4 Arranque em Produção – Novos desafios
- 5 Abordagens alternativas em construção: BUPi@Lab
- 6 Resultados & próximos passos

1.

ALTRAN

- O que fazemos?



LÍDER EM SERVIÇOS DE ENGENHARIA, R&D E TECNOLOGIA



Automóvel



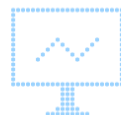
Aeronáutica



Espaço, Defesa e
Naval



Ferrovia, Transportes
e Infraestruturas



Serviços financeiros
e setor público



Energia



Industrial e
eletrónica



Ciência e
saúde



Telecom
e Media



ALTRAN EM NÚMEROS

RECEITA

€ 2.120m em 2016

COLABORADORES

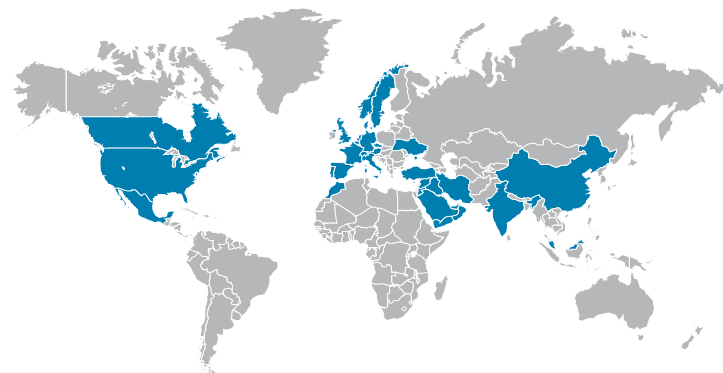
+ 30.000 em 2016

INOVAÇÃO

+ 30 anos

INTERNACIONAL

+ 20 países



- Denmark
- Germany
- France
- Italy
- Scandinavia
- UK
- Spain
- Austria
- Belgium
- Luxembourg
- Netherlands
- Czech Republic
- Switzerland
- Sweden
- Portugal
- US
- China
- India
- Morocco
- Ukraine
- Malaysia
- Middle East
- Canada
- Mexico

ALTRAN MANTEM UMA PRÁTICA GLOBAL DE BIG DATA ANALYTICS

Centros de excelência mundiais com a missão de fornecer serviços globais e especializados de big data analytics

A Altran é reconhecida pela capacidade de executar projetos de **Big Data Analytics globalmente** da Europa para o Mundo.

10
EXPERTISE CENTRES
ANALYTICS

+700
PROFISSIONAIS
BIG DATA
ANALYTICS

+200
PhD DATA
SCIENTISTS

2
WORLD CLASS
CENTERS

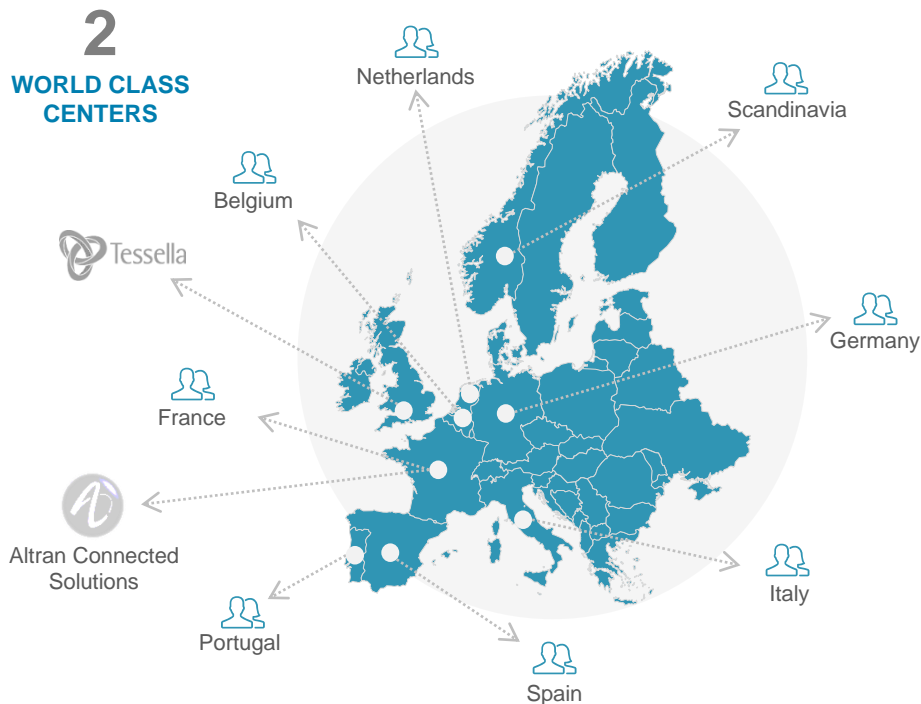
Data engineering | Data fusion and mining | Cognitive Analysis

Bayesian Networks | Time Series | Natural Language | SVMs

Data visualizations | Data optimization | Predictive models

Neural Networks | Machine Learning | Data classification

Mutation | Nonlinear modeling | Clustering applications



2.

O que é o BUPi?

O que é o BUPi? – a missão

Conhecer o território Português,
na medida em que:

- Há propriedades com proprietários desconhecidos
- Imprecisão sobre a sua exata localização e os seus limites

Para quê?

- Gestão e ordenamento do território mais justo e eficaz
- Valorização económica do mesmo
- Prevenção e combate aos incêndios

Projeto-piloto nos seguintes 10 municípios:

Alfândega da Fé | Caminha | Castanheira de Pêra | Figueiró dos Vinhos | Góis | Pampilhosa da Serra | Pedrógão Grande | Penela | Proença-a-Nova | Sertão

O que é o BUPi? – fases do projeto

As diferentes fases do projeto



O que é o BUPi?



Plataforma:
<http://bupi.gov.pt>



Solução digital



Balcão físico

- Processo mais ágil e menos oneroso do que outras iniciativas anteriores
- Processo mais simples e centralizado numa única plataforma digital

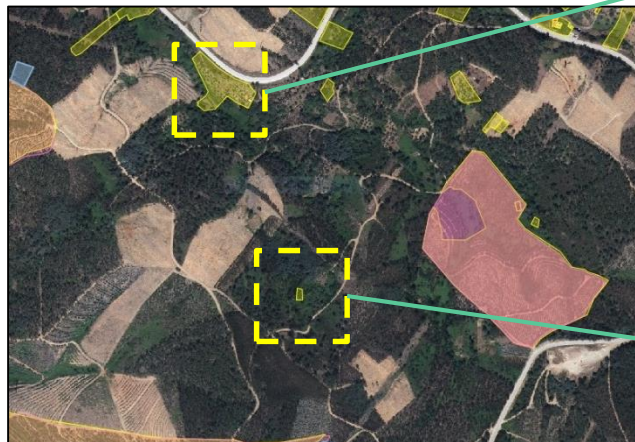
O que é o BUPi? – o processo

Na prática, tem o propósito de:

Identificação das propriedades
através da georreferenciação



Representação gráfica
georreferenciada (RGG)



O que é o BUPi? – o processo

Na prática, tem o propósito de:

Identificação das propriedades
através da georreferenciação



Representação gráfica
georreferenciada (RGG)



Promoção RGG



Registo RGG no
BUPi



Prédio

Matriz

Descrição



Informação de RGG
submetida anotada ao
registo

3.

Dados & Desafios iniciais

Dados & Desafios iniciais – fase As Is

Primeira fase de abordagem aos dados existentes & Mapeamentos

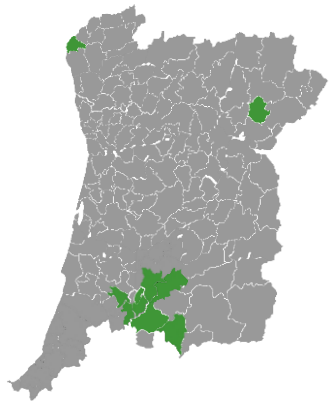


FASE AS IS

A fase AS IS serviu, numa primeira análise, fazer o reconhecimento do território nacional com base em diversas fontes de informação. Foram produzidos Dashboards como demonstradores do panorama nacional e retiradas ilações através dos indicadores produzidos.

As maiores fontes de dados são a AT e o IRN.

Reconhecimento municípios piloto através dos dados



- Alfândega da Fé
- Caminha
- Castanheira de Pêra
- Figueiró dos Vinhos
- Góis
- Pampilhosa da Serra
- Pedrógão Grande
- Penela
- Proença-a-Nova
- Sertão

Nº MATRIZES AT	% MATRIZES AT LOCALIZADAS POR LUGAR/LOCALIZAÇÃO	Nº PRÉDIOS IRN
677.865	% 3	261.735

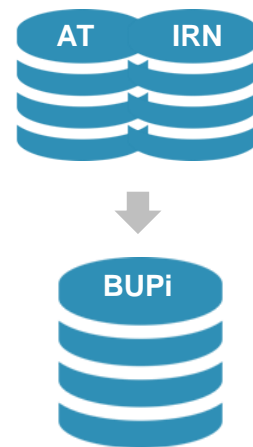
Dados & Desafios iniciais – mapeamentos

Primeira fase de abordagem aos dados existentes & Mapeamentos



MAPEAMENTOS

Por forma a retirar mais valor dos dados fornecidos pela AT e pelo IRN, procedeu-se ao mapeamento dos dados das duas entidades. Para este efeito, foram criadas regras de cruzamentos de dados com vários níveis de precisão. Estes mapeamentos também são essenciais para alimentar a plataforma do BUpi.



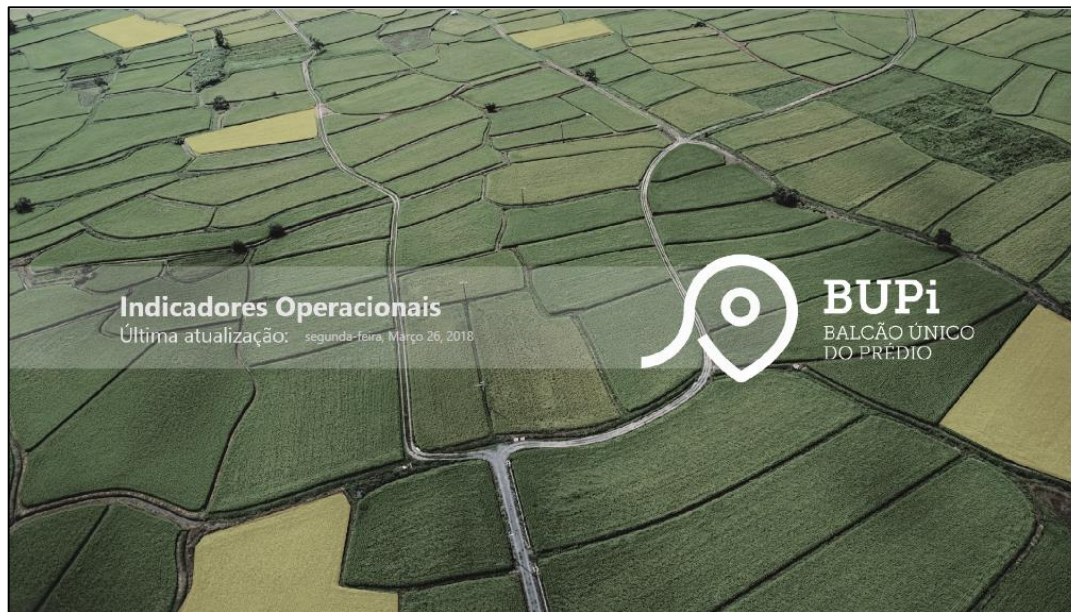
4.

Arranque em Produção – Novos desafios

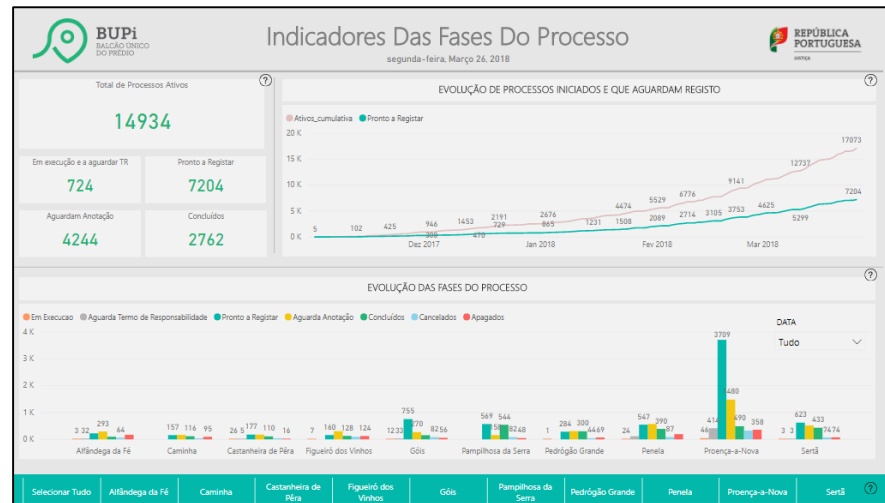
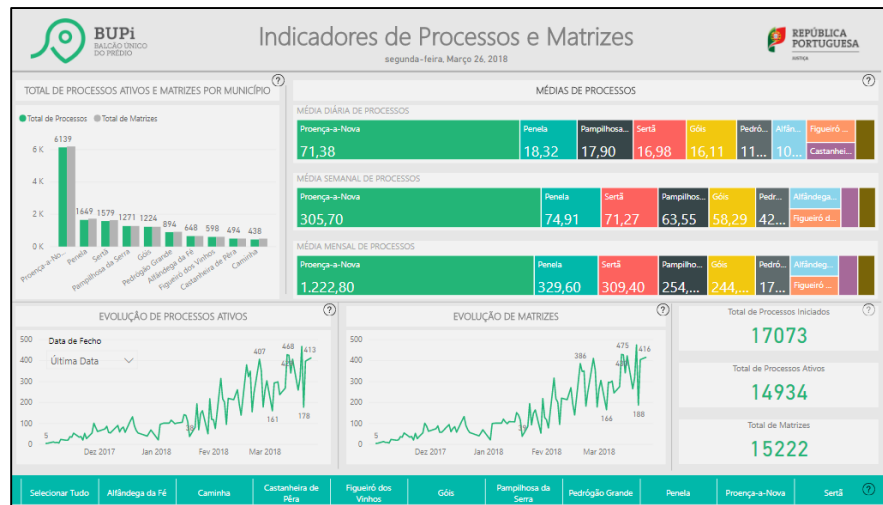
Arranque em Produção – Novos desafios

ANÁLISE OPERACIONAL

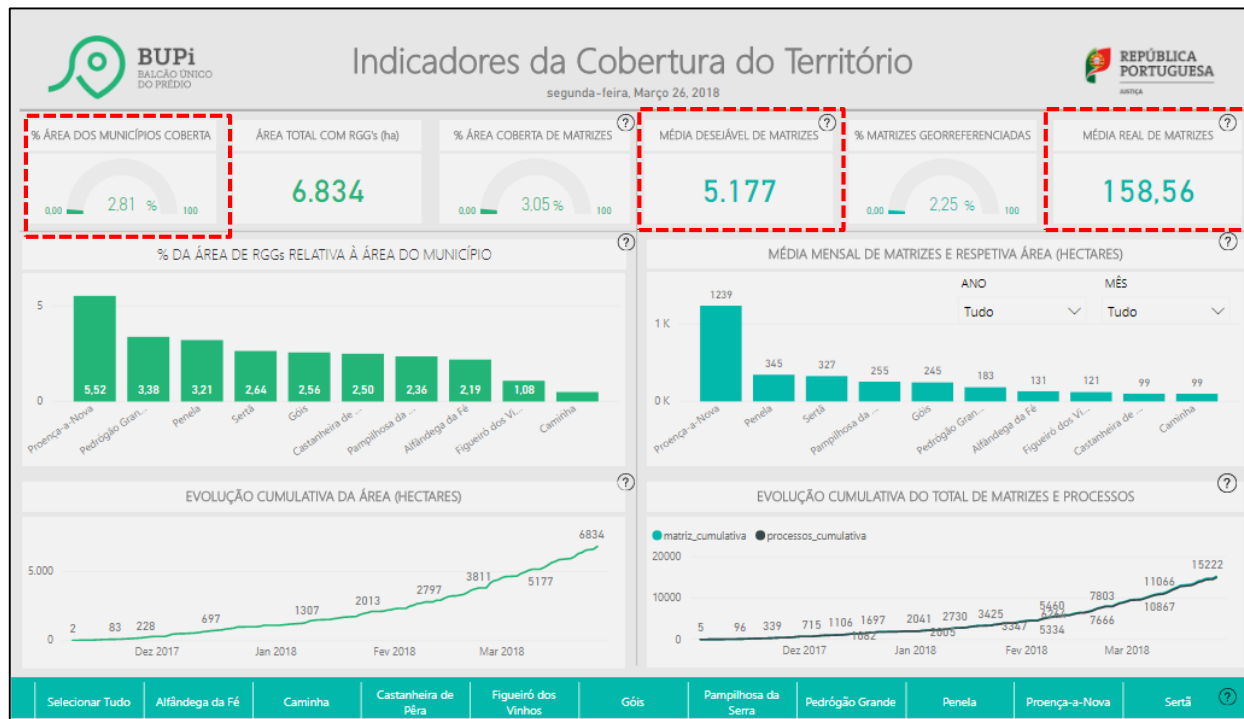
Já com a plataforma do BUPi em produção, é necessário monitorizar a evolução quantitativa dos processos com RGG para se perceber se os objetivos estão a ser cumpridos...



Arranque em Produção – Monitorização



Arranque em Produção – Objetivos



OBJETIVO INICIAL TRAÇADO

100% Municípios cobertos

100% matrizes RGG

OBJETIVO ATUAL

> 50%

5.

Abordagens alternativas em construção –
BUPi@Lab



LAB

Por não estarem a ser atingidos os objetivos de georreferenciação de matrizes, surgiu a ideia de criar uma abordagem alternativa. O Lab serve o propósito de acelerar o processo de conhecimento do território, atuando em vários níveis:



- Localizar uma matriz sem depender do registo do titular
- Acelerar o processo no balcão
- Antecipar necessidade de atuação em certas áreas
- Possibilitar futuro contato personalizado com titulares que não tenham feito o reconhecimento por iniciativa própria

BUPI@Lab – objetivos

- I. Geolocalizar matrizes através das coordenadas geográficas:
 - Por Freguesia 
 - Por Lugar/Localização 
- II. Matching das matrizes com os polígonos naturais através de deteção remota

5.1

BUPI@Lab

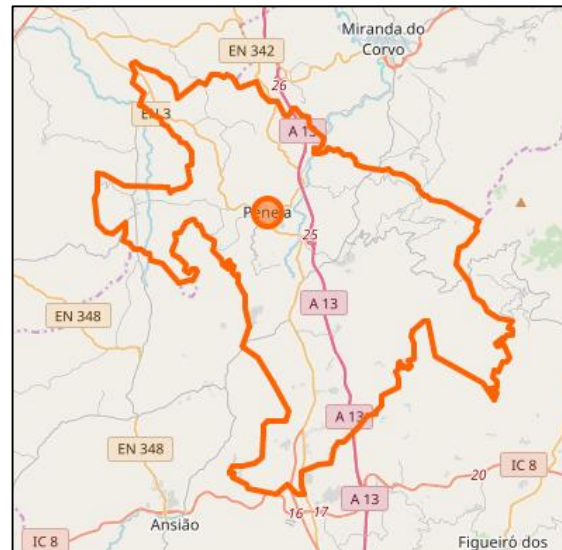
1ª fase: Geolocalização matrizes

BUPI@Lab – objeto de estudo

➤ Município escolhido para o Lab:

Penela

- 70k matrizes AT
- 1.5k RGGs BUPi



BUPI@Lab – pressupostos

Pressupostos:

- Utilização do campo “Localização” (moradas/lugares) da AT
- Mapeamento com dados das toponímias de lugares
- Acesso a dados oficiais de cadastros prediais (polígonos e respectivos números das matrizes)

BUPI@Lab – coordenadas

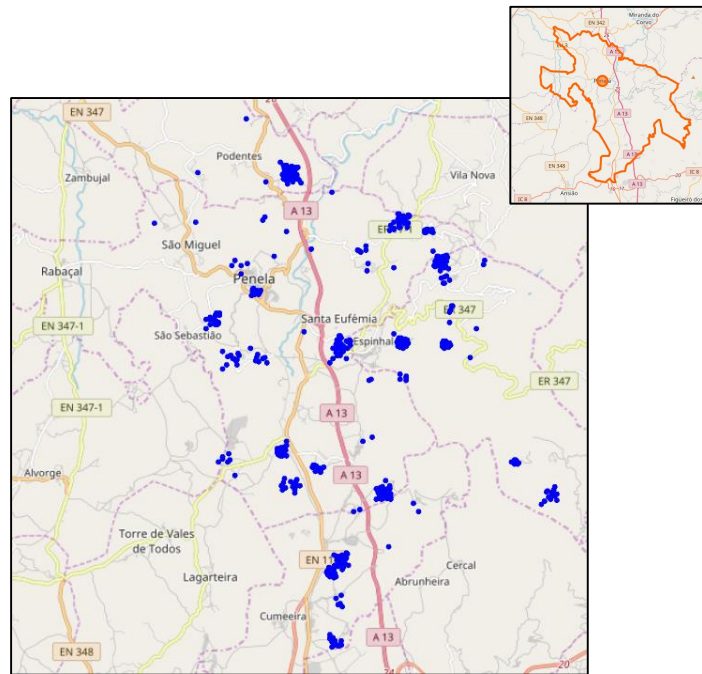
- ✓ Geolocalizar matrizes através das coordenadas geográficas:
 - Por Lugar/Localização

Problema! Não há coordenadas para todos os lugares!

BUPI@Lab – coordenadas

Obtenção de coordenadas geográficas:

- Toponímias oficiais: Câmaras Municipais, entidades (DGT, etc.)
- Toponímias OSM



Problema! Dados escassos ou incompletos!

BUPI@Lab – coordenadas

Solução encontrada para obtenção de coordenadas geográficas:

- BUPi



BUPI@Lab – coordenadas

Solução encontrada para obtenção de coordenadas geográficas:

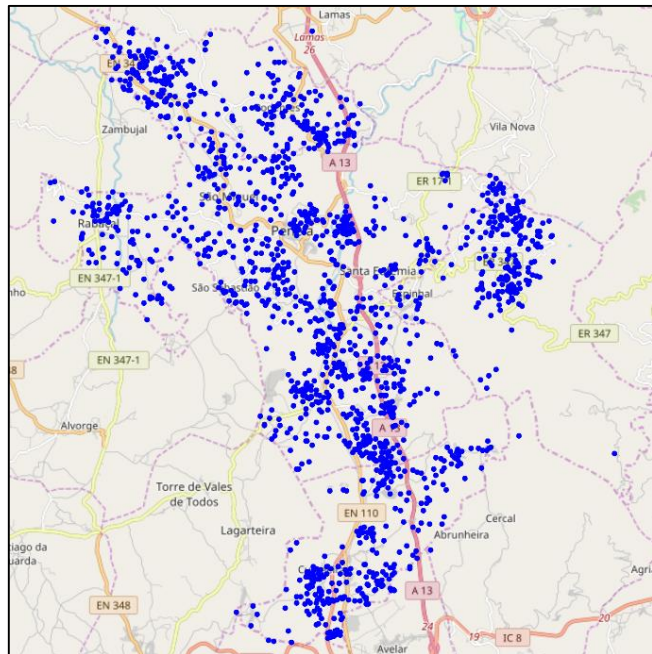
- BUPi



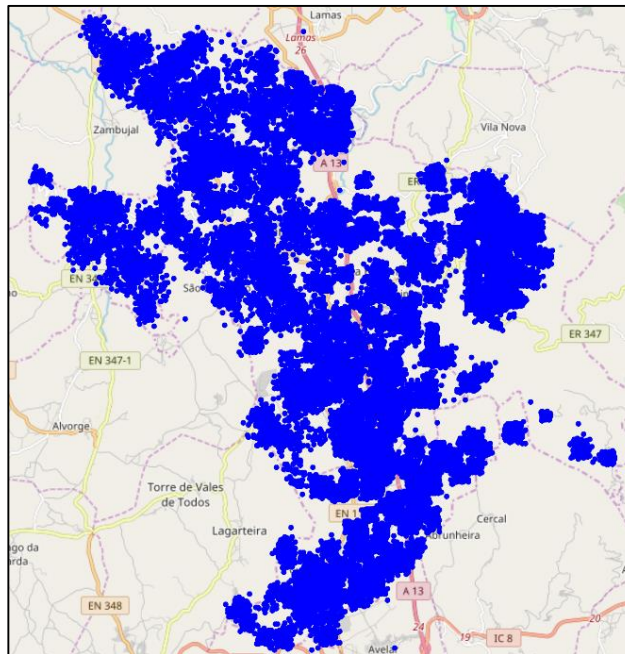
- ✓ Obtendo coordenadas dos polígonos das matrizes: centróides C_M
- ✓ Obtendo coordenadas para os Lugares: centróides medianos C_L

BUPI@Lab – coordenadas

Matrizes BUPi



Matrizes BUPi + geo-BUPi + geo-Toponímia



Impacto:

24x matrizes no mapa
(>36k)

BUPI@Lab – sequenciação numérica

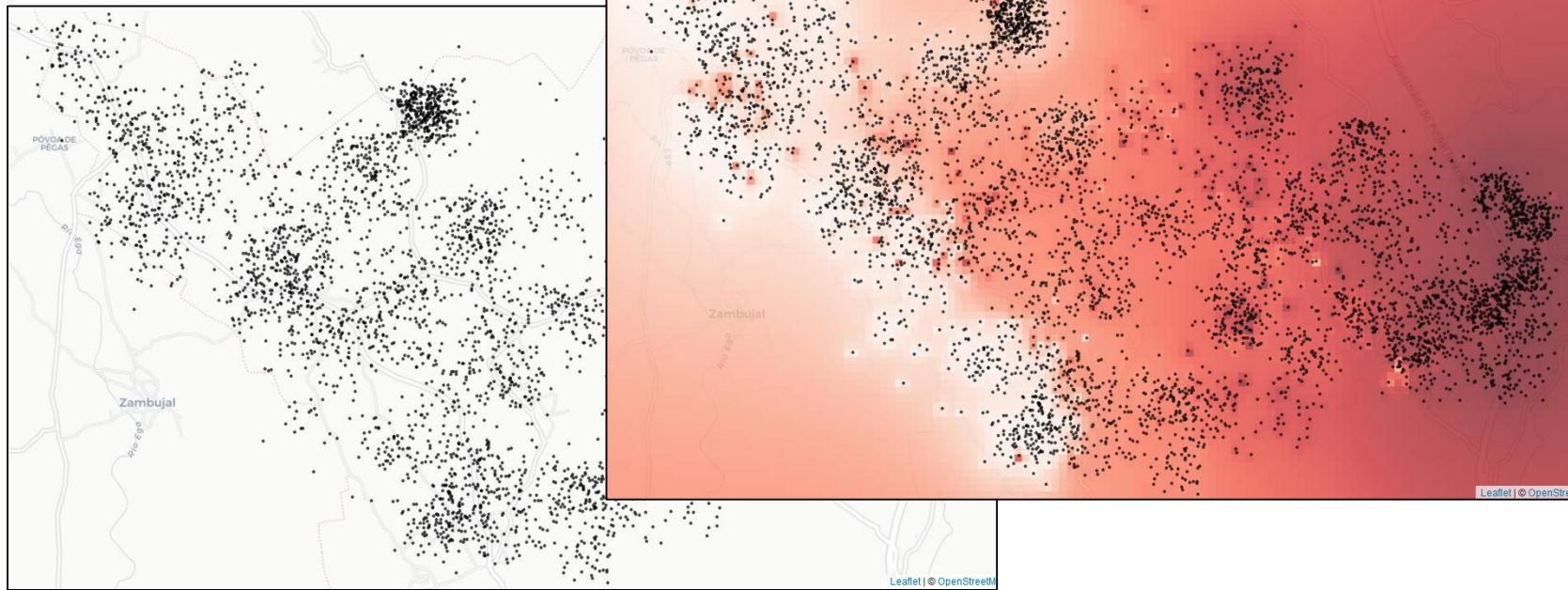
Ainda só conseguimos colocar no mapa cerca de 52% das matrizes.

Existirão outras formas de aproximá-las?

Padrão de distribuição geográfica dos
números das matrizes!

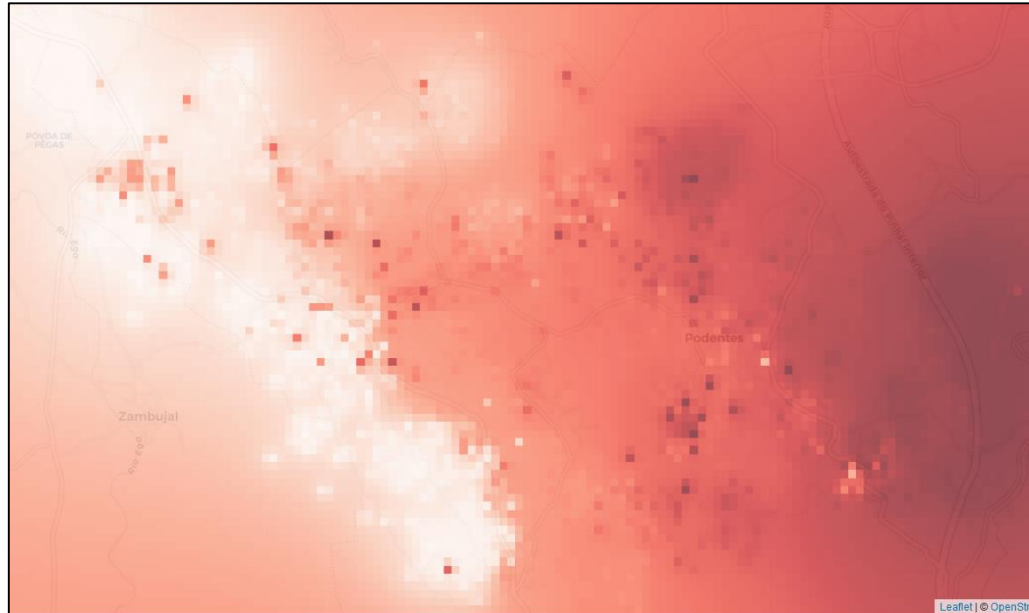
BUPI@Lab – sequenciação numérica

➤ Exemplo com freguesia de Podentes



BUPI@Lab – sequenciação numérica

➤ Exemplo com freguesia de Podentes



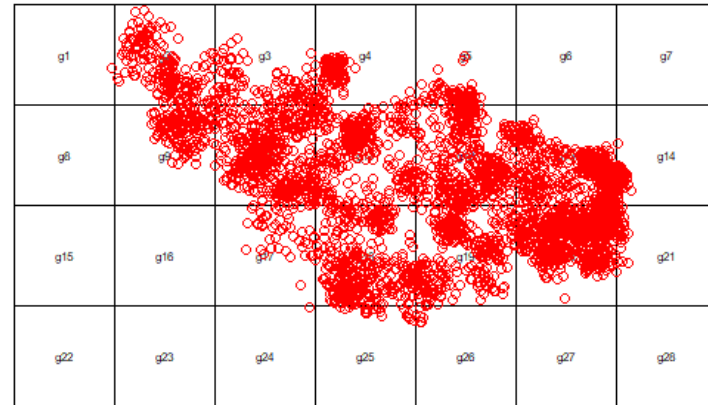
○ Gradiente de Cor:

Claro – número Matriz menor

Escuro – número Matriz maior

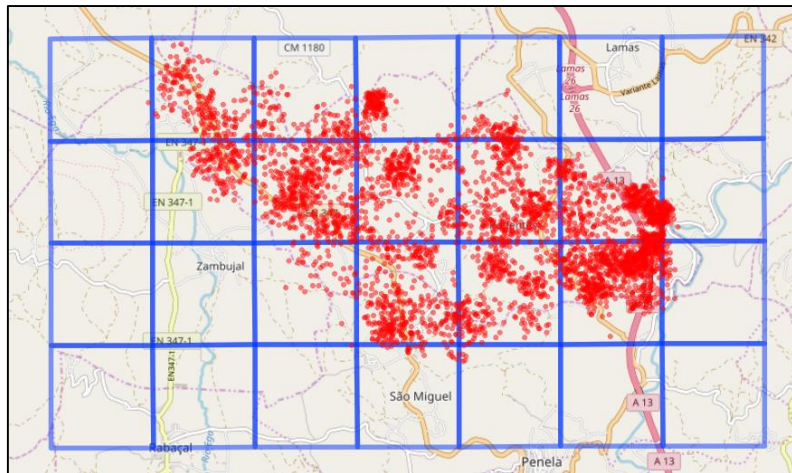
BUPI@Lab – validação aproximação numérica

g1	g2	g3	g4	g5	g6	g7
g8	g9	g10	g11	g12	g13	g14
g15	g16	g17	g18	g19	g20	g21
g22	g23	g24	g25	g26	g27	g28



- Célula: 1500m x 1500m

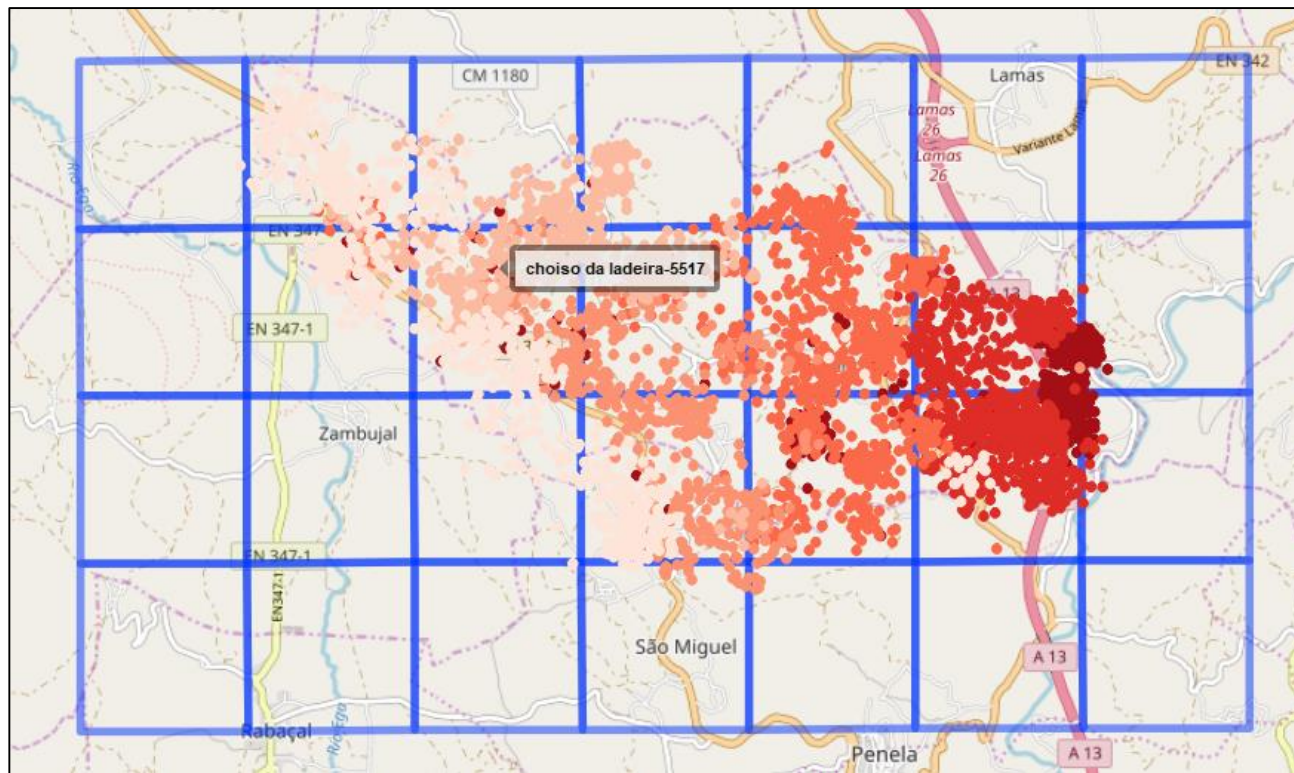
BUPI@Lab – validação aproximação numérica



○ Célula: 1500m x 1500m

- Validação naïve:
 - Samples: 50% referência / 50% validação
 - 80% casos corretos (célula correta)
- Distância mediana:
 - casos corretos: 210m
 - casos incorretos: 340m
- Distância máxima:
 - casos corretos: 1200m
 - casos incorretos: 5400m

BUPI@Lab – validação aproximação numérica



○ Gradiente de Cor:

Claro – número Matriz menor

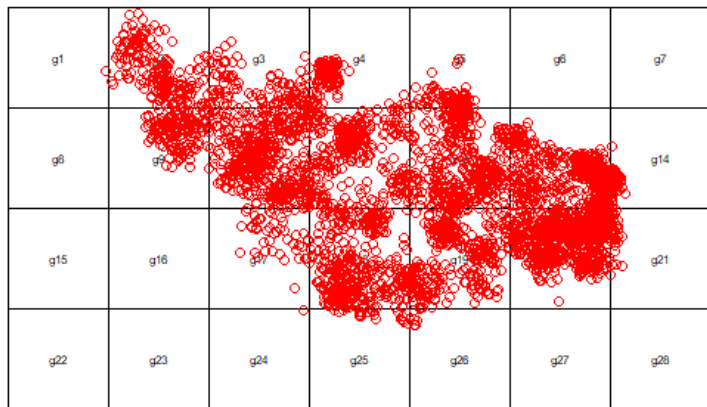
Escuro – número Matriz maior

BUPI@Lab – validação aproximação numérica

Pressupostos:

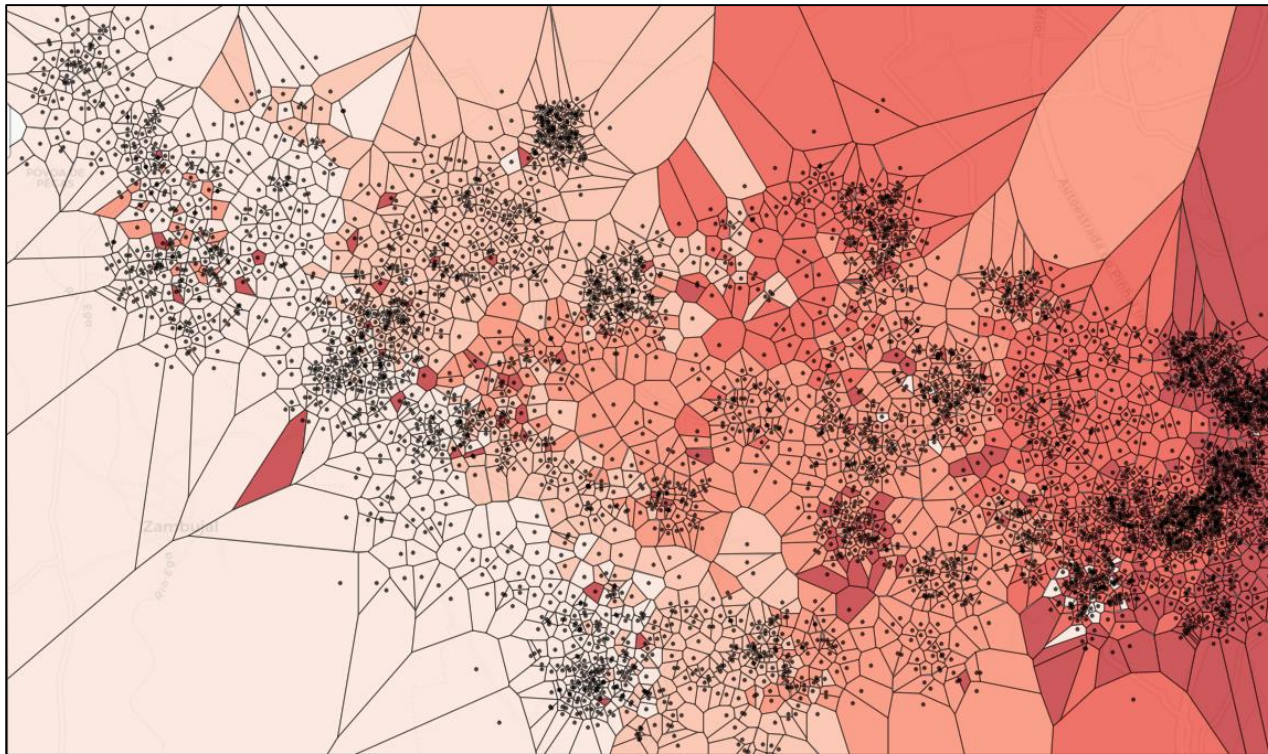
- Padrão de sequenciação geográfica dos números de matriz
- Obtenção de uma coordenada da matriz numericamente mais próxima (potencial matriz vizinha); sinalizar potenciais outliers
- Validação por identificação da célula localizada numa grid e pela distância entre os dois pontos (localizado e real)

BUPI@Lab – sequenciação numérica: a decorrer



- Cada célula tem ID
- Cada ID é uma classe
- Classificação:
 - Accuracy: 81% (XGBoost com Cross Validation)

BUPI@Lab – sequenciação numérica: a decorrer



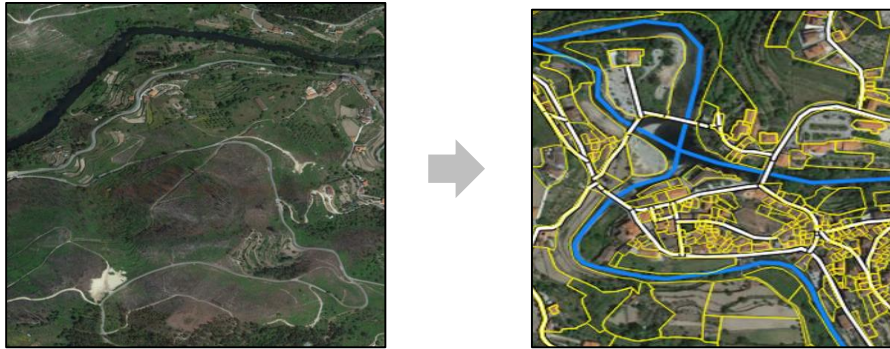
- Polígonos Thiessen/diagrama Voronoi
 - melhorar aproximação numérica
 - usar os centróides dos polígonos para atribuir a coordenada à matriz numericamente mais próxima

5.2

BUPI@Lab

2ª fase: Detecção remota

BUPI@Lab – detecção remota



- Delineação de polígonos através dos limites físicos naturais e/ou artificiais

BUPI@Lab – output Geolocalização + Detecção remota

Geolocalização



Detecção remota



- Matching das matrizes geolocalizadas na 1ª fase com os polígonos delineados através da detecção remota

6

Resultados & próximos passos

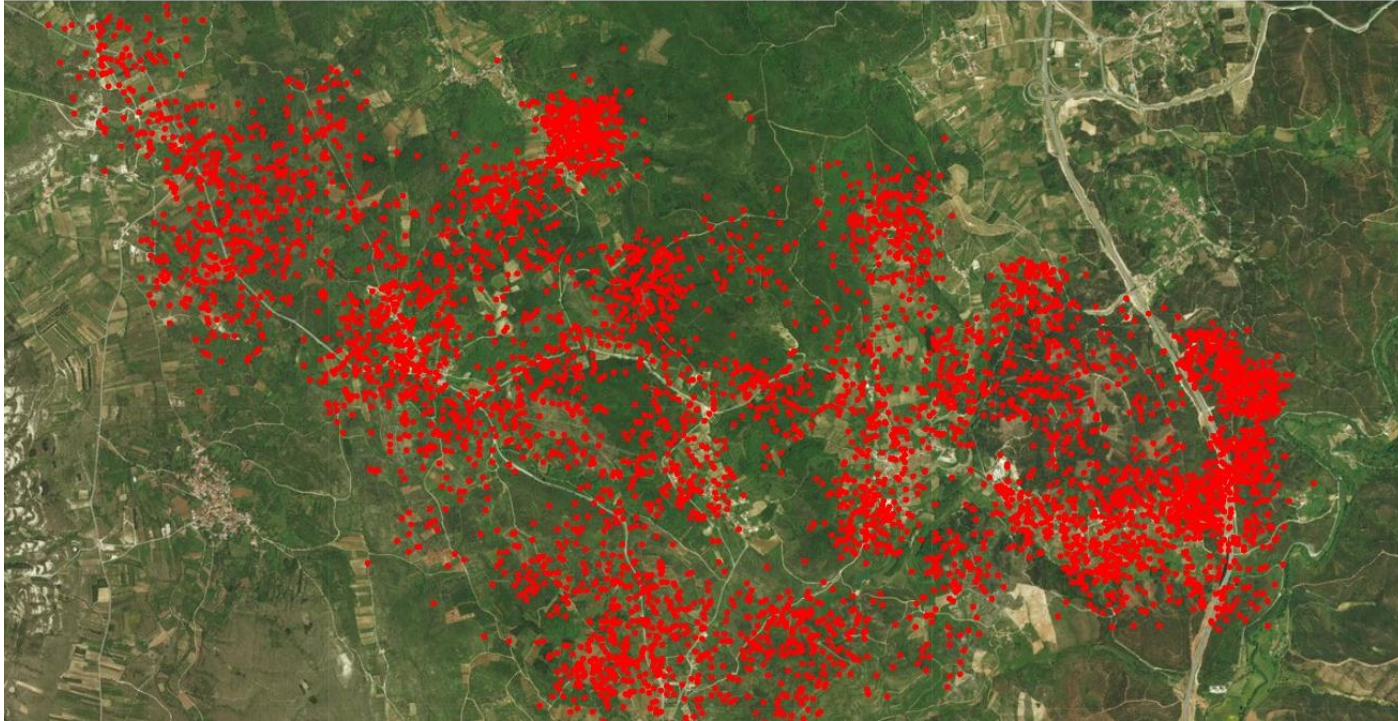
BUPI@Lab – Resultados: Freguesia Podentes

Matrizes BUPI

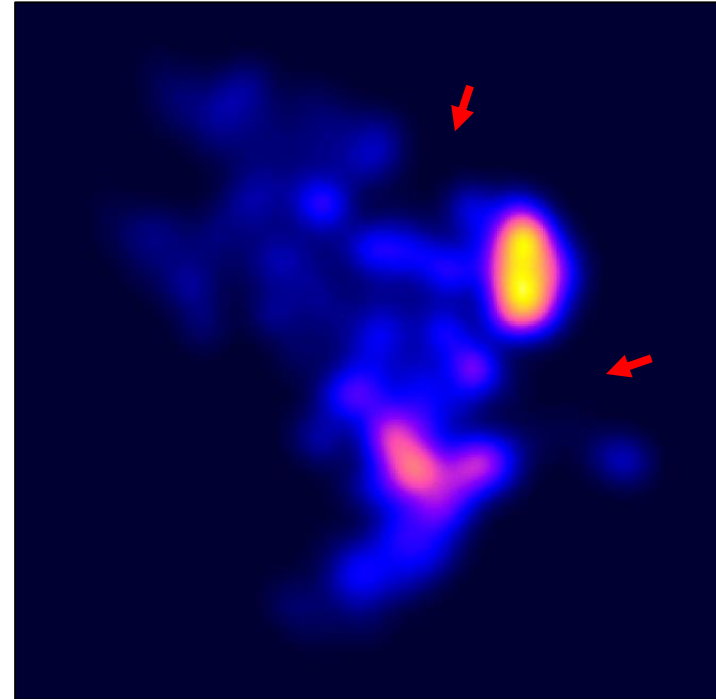


BUPI@Lab – Resultados: Freguesia Podentes




Matrizes AT



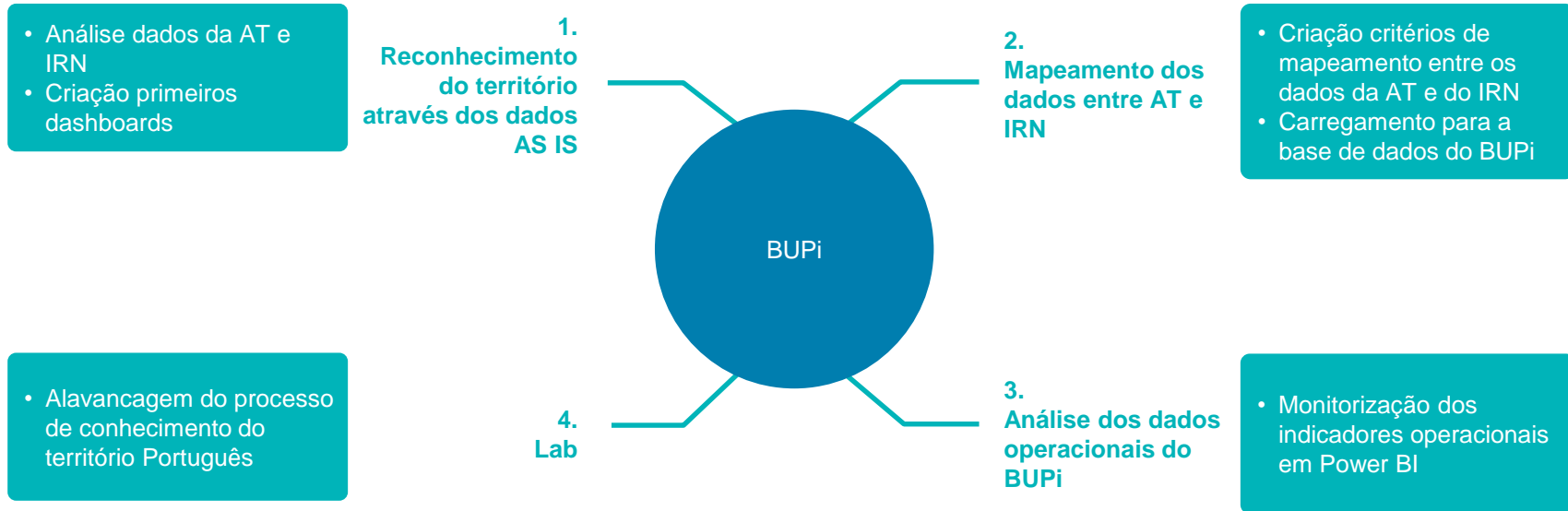
BUPI@Lab – Resultados: Penela



BUPI@Lab – Resultados / ganhos imediatos

- Globalmente:
 - Território conhecido > 50% 
 - Maior precisão sobre a localização das terras com dono conhecido e terras sem dono conhecido 
- A implementar em produção:
 - operacionalmente, redução do tempo de criação de uma RGG 

Overview Altran@BUPi



BUPI@Lab – próximos passos?

- Usar polígonos Thiessen para melhor aproximar matrizes
 - Continuar a trabalhar num bom algoritmo de classificação
 - Análise do campo das *confrontações* para maior precisão: graph networks e outros algoritmos que envolvam análise de texto
 - Detecção remota [dependência]
 - Usar centróides dos polígonos da detecção remota e prever número matriz (interpolação, decision trees, etc.)
- Estender aos restantes Municípios do projeto-piloto

Obrigado!

Contactos:

Daniel Andrade
Data Scientist / Researcher
daniel.andrade@altran.com

Jorge Afonso
Head of Big Data Analytics Expertise Centre
jorge.afonso@altran.com

alTRan