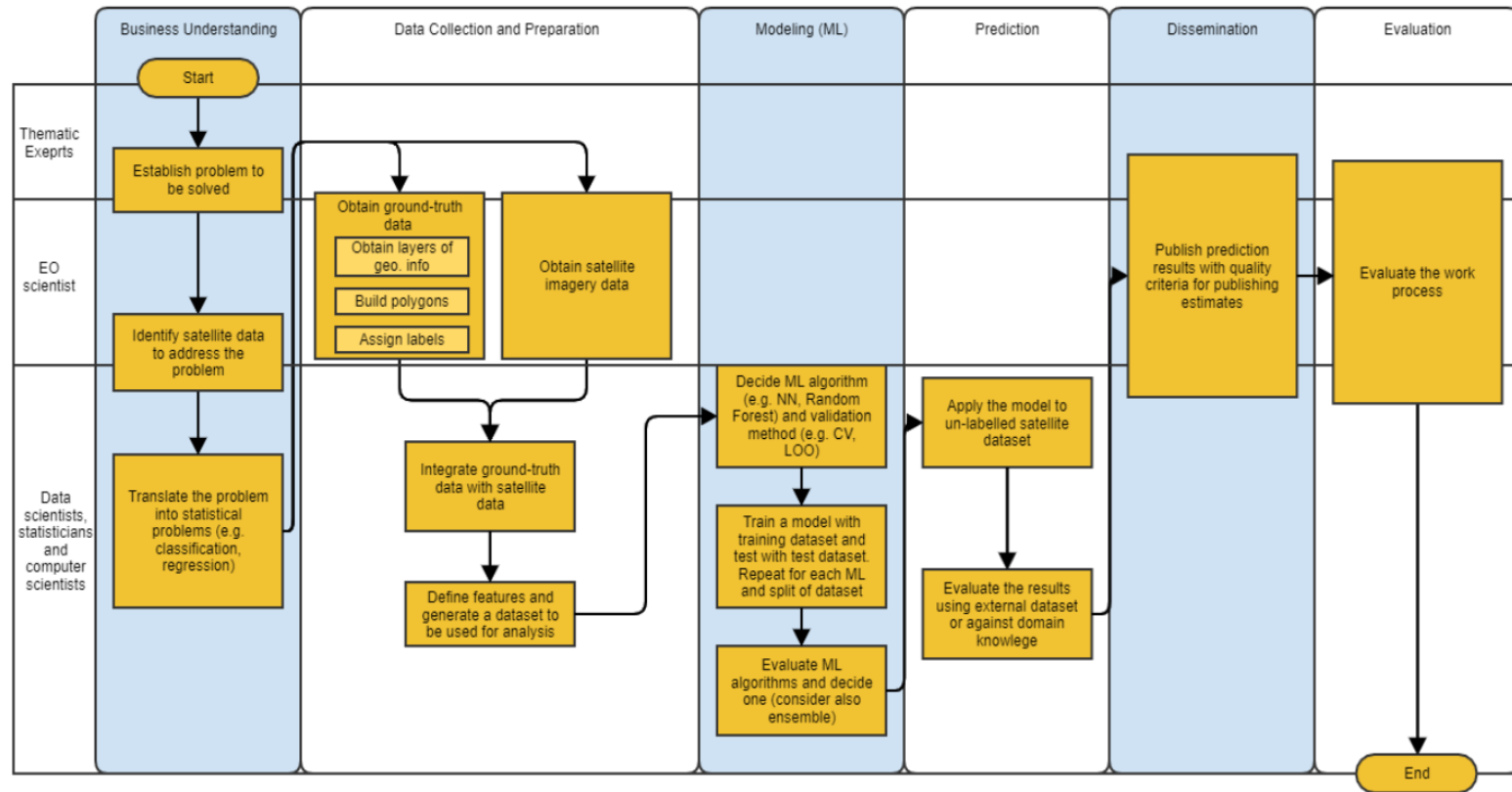IMAGERY THEME REPORT

ABEL CORONADO
JIMENA JUÁREZ

STRUCTURE OF THEME REPORT

- Introduction to the report

- Generic Imagery Pipeline

- Key aspects & Questions regarding pilot projects

IMAGERY PIPELINE

KEY ASPECTS & QUESTIONS
REGARDING PILOT PROJECTS

- Background

- Organisational context

- Data context

- ML Solutions

- Results, outcomes and going beyond

BACKGROUND

- What is the problem to solve?

- Is the project's motivation derived from the need to reduce costs (time and money) or to make up for unavailable/unreliable national datasets ?

- Is the project an improvement or the creation of a new process? If it is an improvement, how is it assessed?

COUNTRY	PROBLEM TO SOLVE	CONTRIBUTION	VALUE ASSESSMENT
Australia	Reduce the amount of manual intervention required during regular Address Register (AR) maintenance processes.	Reduce costs (time) by improving the current process that is a resource intensive process.	The number of automatically classified addresses.
Netherlands	Detect poverty and population distribution from aerial or satellite imagery.	Learn how to use machine learning to exploit imagery as a new data source in the production of official statistics	A working computer prototype.
Switzerland	Facilitate land use and cover classification and by improving change detection	Improve existing process to reduce costs (time).	A working computer prototype that allows to demonstrate the innovative potential of the FSO in the use of artificial intelligence to process images.
Mexico	Detect the extension of urban areas nationwide using ML	Reduce time and money. Generate information products that contribute to the cartographic update.	Clear objectives with links to potential impacts on existing and future data products.

ORGANISATIONAL CONTEXT

- What is the relevance of this problem for the objectives of your Institution?

- Who are the stake-holders involved?

- What is the relevance for the main stake-holders?

- Are there expectations to go beyond the demonstration phase? Is there an roadmap?

DATA CONTEXT

- How did you decide to use these datasources? What variables were considered? Were there other datasets available?

- What kinds of preparation processes were applied to the datasets?

Country	Imagery	Number of Labeled Images	Number of Classes Labeled
Australia	Aerial Images	6,000	6
Netherlands	Aerial Images	70,000	5
	Landsat 8	70,000	
Switzerland	Aerial Images	420,000	73
	Landsat 8		
Mexico	Landsat 5,7	40,000	2

ML SOLUTIONS

- What ML solution did you try and why?

- Do you consider your solution common or experimental?

- How does this decision relate to the initial motivation or this project (reducing costs or making up for unavailable datasets)?

Country	Algorithm(s) tested	Python Library	Achievement
Australia	• Custom 12 layers CNN Architecture	• Tensorflow (CPU)	Moved to production
Netherlands	• CNN Architecture based on VGG16 and ResNet50 • RF and SVM	• Tensorflow (GPU) • Scikit-learn	Still proof of concept
Switzerland	• CNN Architecture based on Xception • RF	• Tensorflow (GPU) • Scikit-learn	Still fine-tuning the algorithm
Mexico	• CNN Architecture based on LeNet • Extra-Trees	• Tensorflow (CPU) • Scikit-learn	Still proof of concept

RESULTS, OUTCOMES & GOING BEYOND

- To what extent was the initial problem solved?

- If there is an expectation to go beyond a demonstration phase, how far down the roadmap is the project?

- What were the main challenges faced?

- What were the main gains obtained?

- What do you consider the added value is?

RESULTS, OUTCOMES & GOING BEYOND

LESSONS LEARNED

- Have a solid business case for the ML project

- Narrow down the problem; just enough complexity to add value.

- DNN training involves the use of specialized hardware for confidential data it is required to incorporate this hardware into the internal computer centers.

- To be able to carry out classification exercises using imagery: it is essential to have high quality training sets validated by thematic experts, field work as well as complementary data sets.

¡GRACIAS!