

# living planet symposium | BONN 23-27 May 2022

TAKING THE PULSE  
OF OUR PLANET FROM SPACE



## Federated Learning for Vessel Detection with AIS as Training Data

Gonçalo Sousa Mendes

27 May 2022



# Privacy Preserving Machine Learning Project

Goal: Use machine learning to harness the power of EO data, while protecting sensitive data sets

start: May 2021

end: Jan 2023

1. Land applications:
  - a. Socio economic mapping
2. Ocean applications:
  - a. **vessel detection**
  - b. bycatch prediction

## Partners



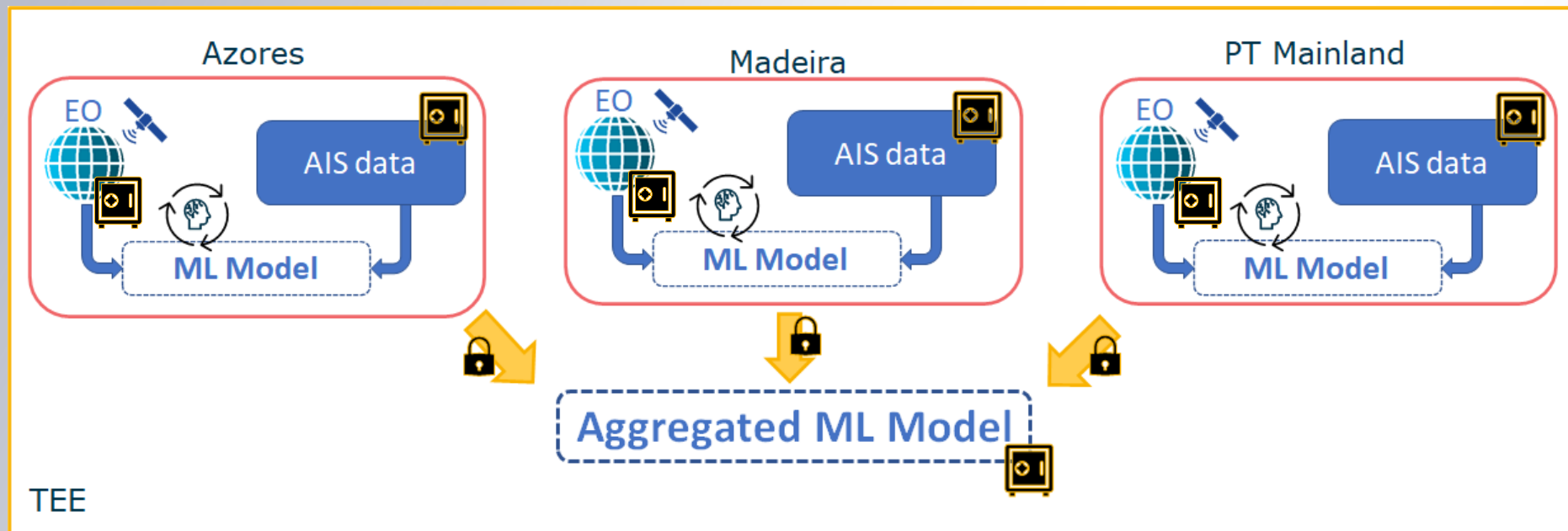
This project has received funding from the European Space Agency Contract No. 4000134424/21/I-NB.

# Context and Motivation

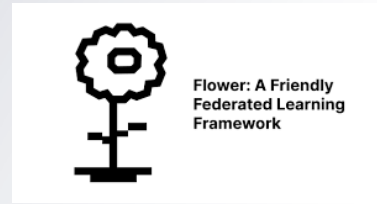
- Fishing activity stressing biodiversity -> Laws: VMS, **AIS**, e-logbooks
- Location of fishing vessels might indicate potential good fishing grounds
- Susceptible to fraud and misguidance

# Solution

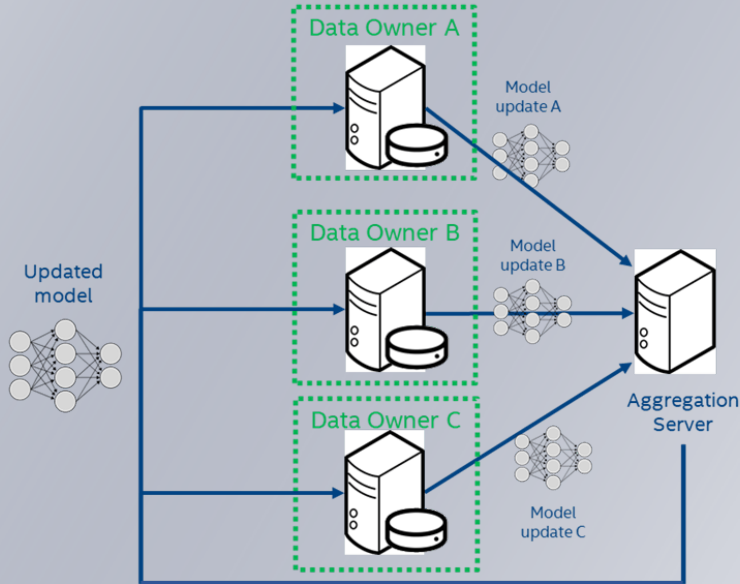
- AIS + Sentinel2L1C: Monitor Illegal, Unregulated and Unreported data (IUU):
  - Deep Learning



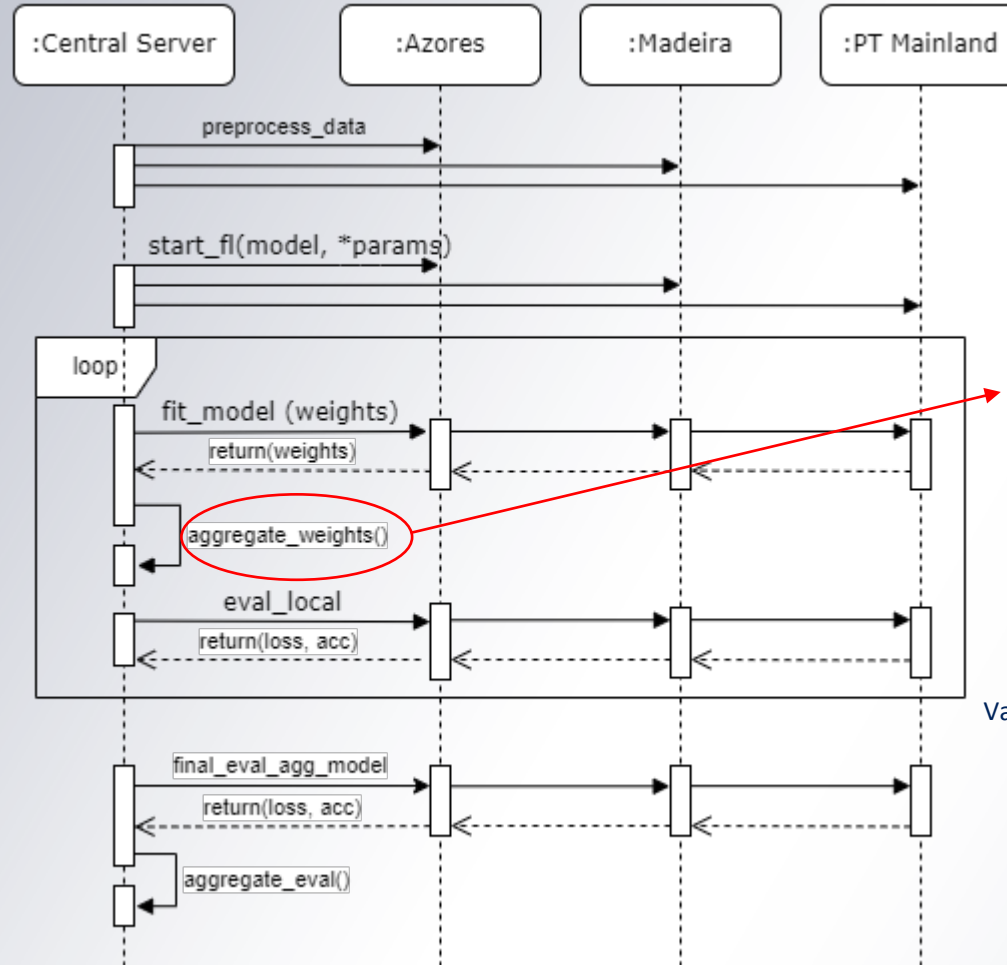
# Federated Learning



Federated Learning Architecture



- Clients contain data that cannot be shared
- Aggregation server coordinates model updates



Weighted average

$$w_{t+1} \leftarrow \sum_{k=1}^K \frac{n_k}{n} w_t^k$$

k: client number  
 t: round  
 n<sub>k</sub>: kth data set size  
 n: total data set size

Vanilla Federated Averaging (McMahan et al., 2017)

# Data

## S2L1C:

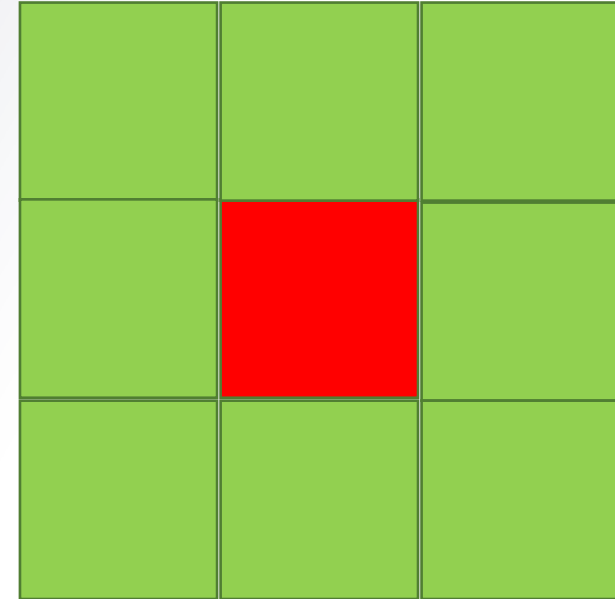
- Open data
- Fully available

## AIS:

- Several millions points
- TOI: 2015 to 2018
- AOI: Portuguese EEZ

# Vessel Detection Approach

- Classification: vessel/not vessel
- Use AIS to create the vessel tile
- Get surrounding tiles as non-vessel data



Binary classification:

- Vessel
- Not Vessel



50 pixels

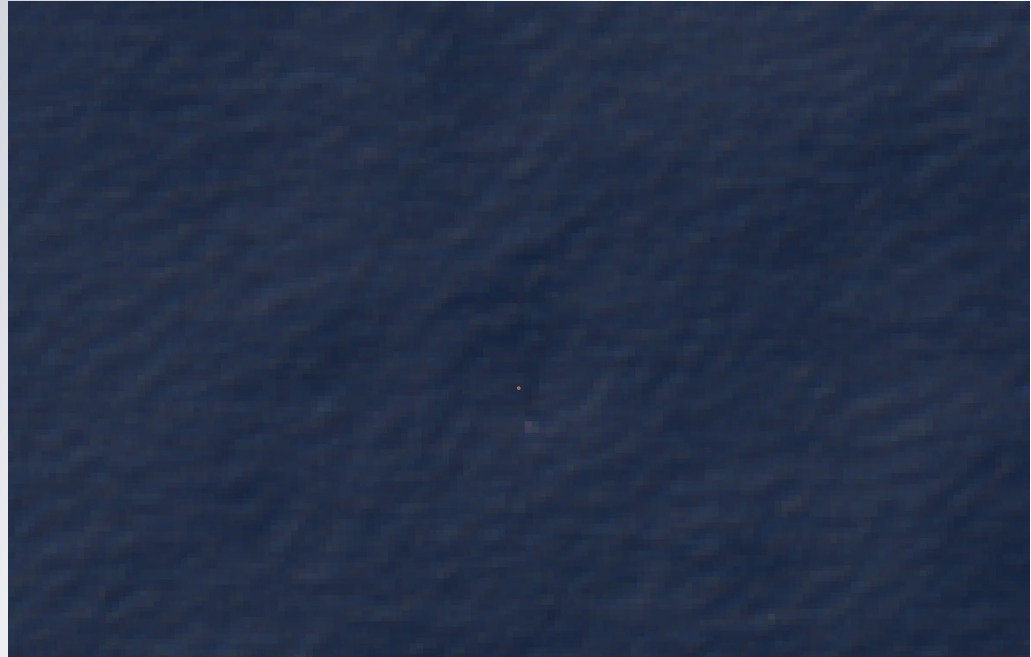
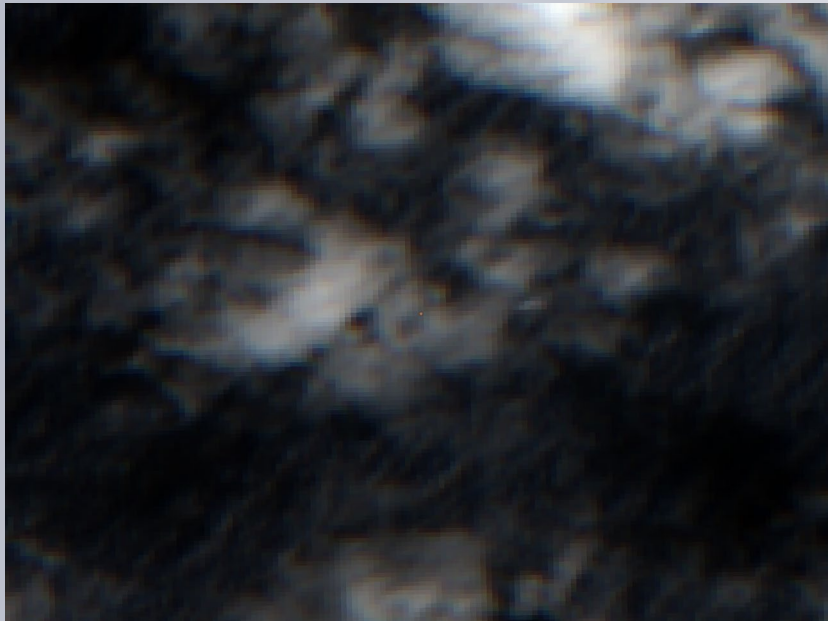


Length Overall x Breadth  
Extreme: **29.25 x 8 m**



# Data

- Ships too small
- Cloud algorithm of S2L1C is no perfect
- Not a lot of vessel images



# Pre-processing data

Data not identically distributed

Azores



5.1 million

Madeira



2 million

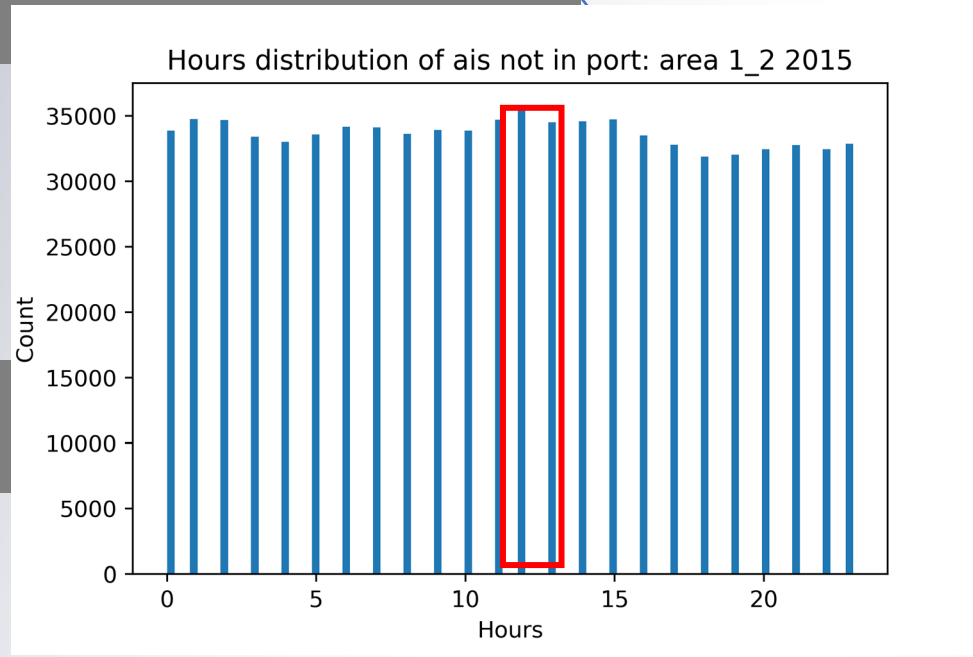
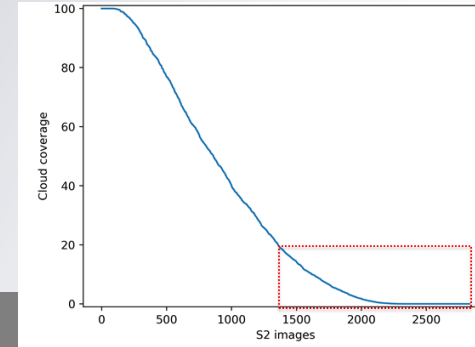
PT Mainland



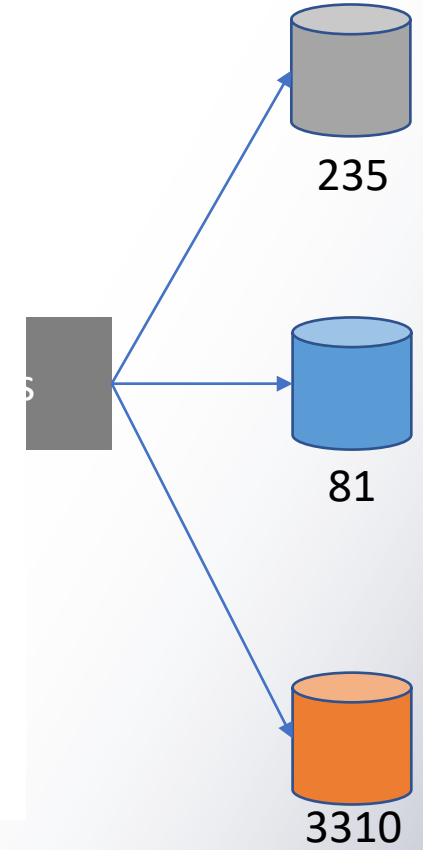
32.3 million

Get S2 metadata for AOI and TOI

Download S2



Very few vessel tiles



# Experimental setup

- Compare:
  - FL model
  - All data model
  - Individual models
- Rounds as parameter
- Epochs as parameter
- Epochs for individual models:
  - rounds \* epochs in FL

Model:  
Same class for  
all instances

Machine:  
50 cores  
128 Gb RAM



# Next Steps

- Test models
  - Epochs vs Rounds
  - FL strategies
  - Segmentation approaches
  
- Add TEE to the setup

# Conclusions

- Several good tools for Federated Learning (Flower, IBM FL, OpenMined PySyft, TensorFlow FL, etc.)
- EO data needs to be downloaded:
  - Goes against trend of putting app in the data source
  - **Download is expensive (time and resource)**
  - Using API may expose fishing grounds

# Conclusions

- Assuring that the tile contains a vessel in the middle:
  - It may require access to the data to control the quality
  - Not seeing the tiles created makes it difficult to assess the false negatives



# Thank you

## References

- McMahan, B., Moore, E., Ramage, D., Hampson, S., and y Arcas, B. A. Communication-efficient learning of deep networks from decentralized data. In Singh, A. and Zhu, X. J. (eds.), Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, AISTATS 2017, 20-22 April 2017, Fort Lauderdale, FL, USA, volume 54 of Proceedings of Machine Learning Research, pp. 1273–1282. PMLR, 2017. URL <http://proceedings.mlr.press/v54/mcmahan17a.html>.

Check info here:



Visit us at **Booth #6**