Plano de gestão de dados dinâmico – a experiência do projeto SAIL

Susana Barbosa, Yulia Karimova, Cristina Ribeiro

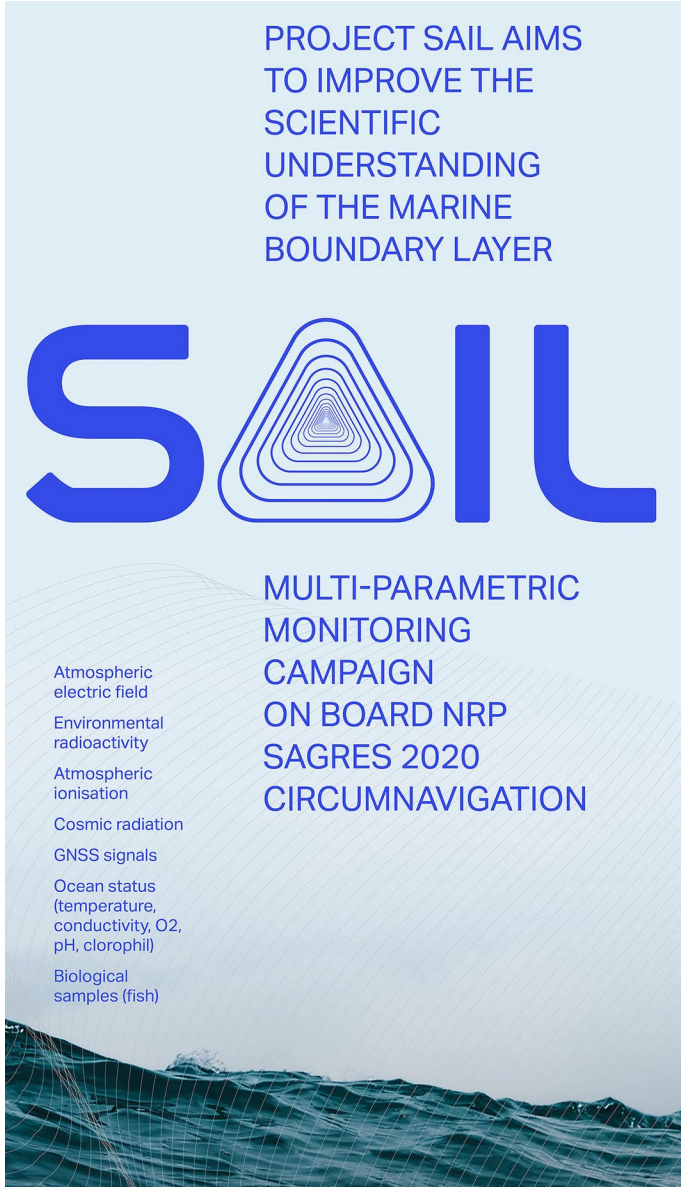
INESC TEC

SAIL

Space-Atmosphere-Ocean Interactions in the Marine Boundary Layer

Projeto SAIL

Campanha de monitorização da camada limite marinha a bordo do navio-escola NRP Sagres



Campanha de monitorização

Janeiro – Maio 2020

Follow-up

Maio 2020 - presente

+15 parâmetros geofísicos

+10 GB / dia

+ 500 dias

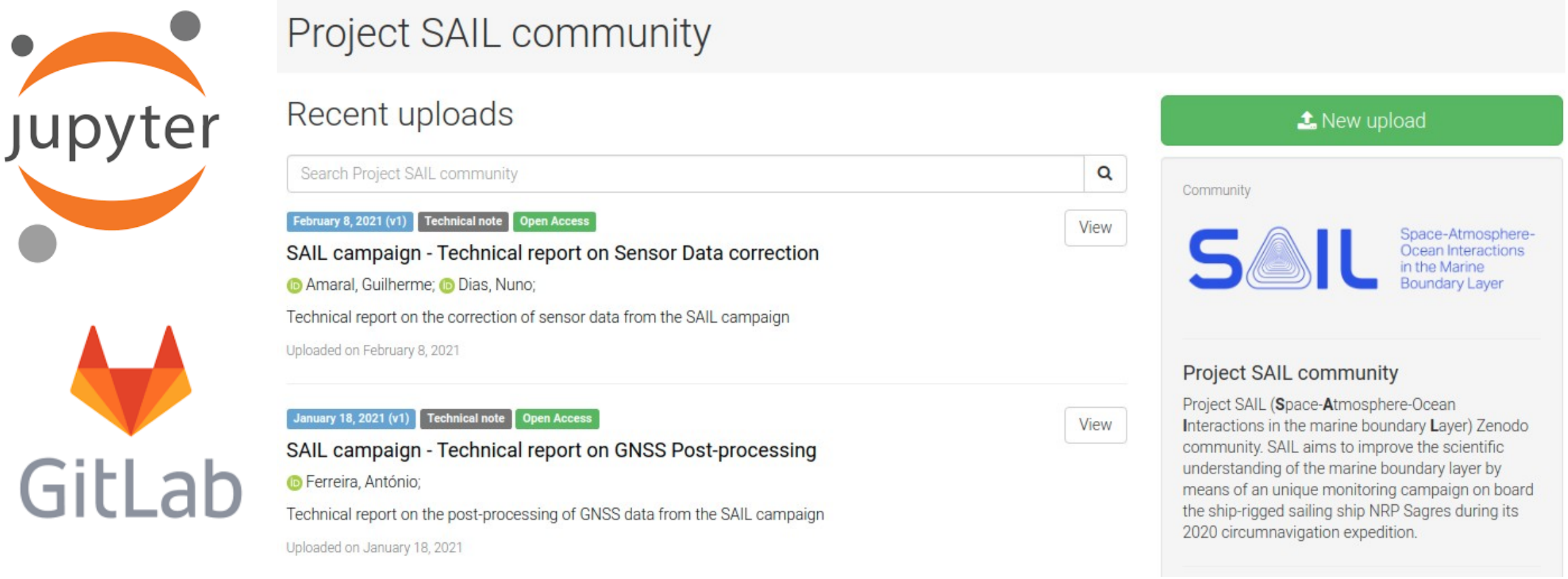
Projeto SAIL

Interdisciplinaridade

Ciência aberta

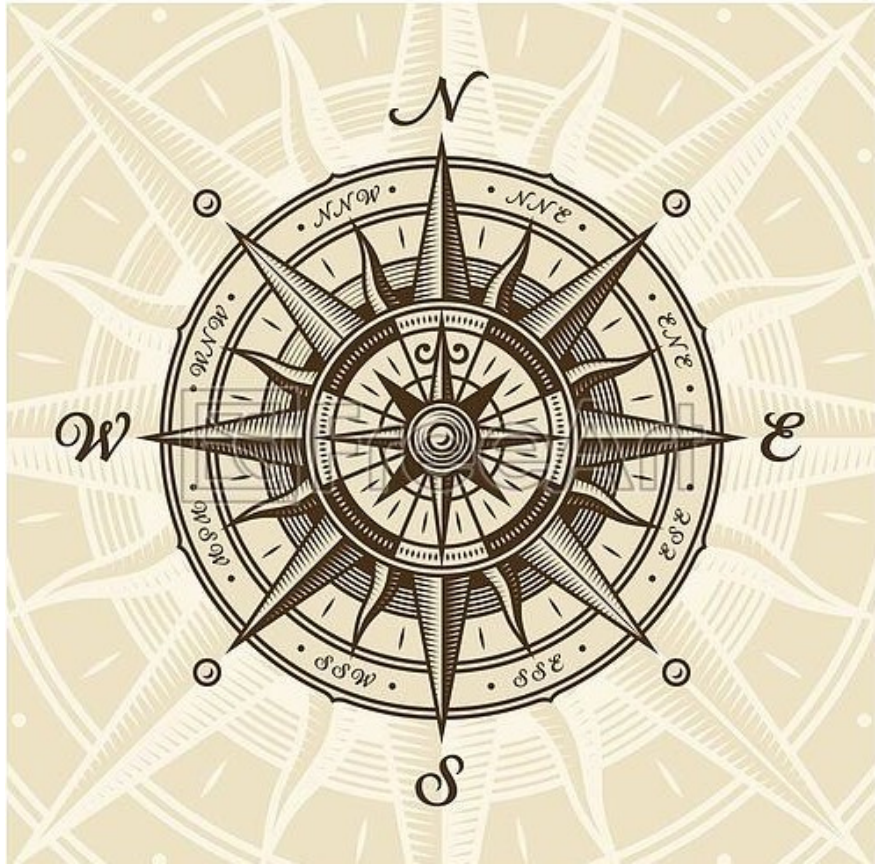
Projeto SAIL

Reproducibildade (dados, meta-informação, código & software)



Plano de gestão de dados do projeto SAIL

Motivação é fundamental!



Plano de gestão de dados do projeto SAIL

Versões “oficiais”

Fevereiro 2020

Abril 2020

Junho 2020

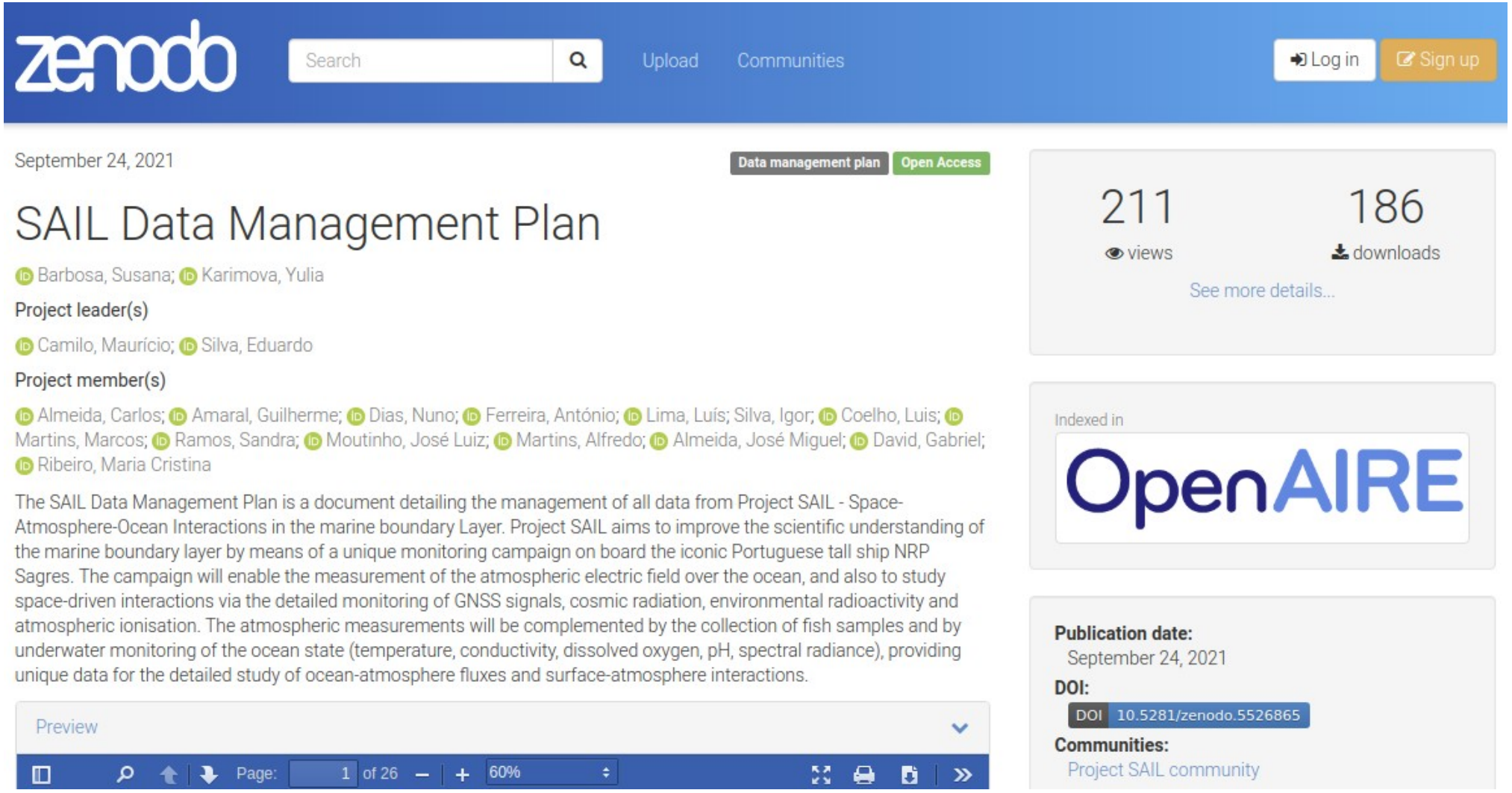
* Novembro 2020

* Março 2021

* Setembro 2021

* Zenodo

Plano de gestão de dados do projeto SAIL



Plano de gestão de dados do projeto SAIL

Nível de detalhe - exemplos

Table 4 summarizes the naming conventions adopted for all the data from the project, which are formed as: type of activity + group of data infix + measured parameter infix.

	Testing	Circumnavigation	Sagres
Ship data	TEST_SHIP_*	SAIL_SHIP_*	SAGRES_SHIP_*
Sensor data	-	SAIL_SD_*	SAGRES_SD_*
Geosensor data	-	SAIL_GD_*	SAGRES_GD_*
Processed data	-	SAIL_PD_*	SAGRES_PD_*

Table 4: dataset naming conventions; * denotes the parameter infix described in Table 2.

Plano de gestão de dados do projeto SAIL

Nível de detalhe - exemplos

compressed files name	individual files
*_SHIP_CI_yyyymmdd.tgz	yyyymmdd-10s.records yyymmdd-1s.records yyyymmdd-block.records yyyymmdd.log
*_SHIP_E1_yyyymmdd.tgz	E1_yyyymmdd_HH.txt
*_SHIP_E2_yyyymmdd.tgz	E2_yyyymmdd_HH.txt
*_SHIP_GA_yyyymmdd.tgz	GA_yyyymmdd_HH.txt
*_SHIP_NS_yyyymmdd.tgz	GPS1_A1_oem4_yyyymmdd_HH GPS1_A2_oem4_yyyymmdd_HH GPS1_IMU_yyyymmdd_HH GPS1_NMEA_yyyymmdd_HH GPS2_A1_oem4_yyyymmdd_HH GPS2_A2_oem4_yyyymmdd_HH GPS2_IMU_yyyymmdd_HH GPS2_NMEA_yyyymmdd_HH GPS3_A1_oem4_yyyymmdd_HH GPS3_A2_oem4_yyyymmdd_HH

Plano de gestão de dados do projeto SAIL

Nível de detalhe - exemplos

compressed files name	individual files
*_SHIP_CI_yyyymmdd.tgz	yyyymmdd-10s.records yyymmdd-1s.records yyyymmdd-block.records yyyymmdd.log
*_SHIP_E1_yyyymmdd.tgz	E1_yyyymmdd_HH.txt
*_SHIP_E2_yyyymmdd.tgz	The electric field hourly datafiles E1_yyyymmdd_HH.txt (upper mast sensor) contained in the compressed file *_SHIP_E1_yyyymmdd.tgz have the following structure: col 1: timestamp (seconds.microseconds) col 2: date (mm/dd/yyyy) col 3: time (HH:MM:SS) col 4: voltage (power) (V) col 5: voltage (internal) (V) col 6: Panel temperature (deg C) col 7: Electric field (V/m) col 8: Leakage current (nA) col 9: CS110 status (numeric code) col 10: Internal RH (%) col 11: shortwave incoming radiation (W/m2) col 12: shortwave outgoing radiation (W/m2)
*_SHIP_GA_yyyymmdd.tgz	
*_SHIP_NS_yyyymmdd.tgz	

Notas finais

Plano de gestão de dados do projeto SAIL

- instrumento fundamental

- motivação importa

- detalhe (nunca é demais!)

- “*release soon, release often*”

Obrigada!

@sail_sagres