SOCIB Data Management System

From regional to European impact

Juan Gabriel Fernández Pineda
Madrid, September 19, 2023
Outline

SOCIB in a nutshell

SOCIB Data Repository

Data Management System

Link to EOSC and increased impact
SOCIB in a nutshell
A multiplatform ocean observing & predicting system: from the coast to the open sea, from events to climate.

3 DRIVERS
- Science priorities
- Technological Development
- Society Needs

ACCESS MODALITIES
- Open Data Access
- Permanent Observing Programs
- Open Access to Infrastructures

COLLABORATIVE
- CSIC, IEO, UIB

ASSESSMENT
- Every 4 years

“Be a reference research infrastructure in the field of marine R&D&I and contribute to the development of a digital twin in the Mediterranean to support decision-making in a context of climate change”

TIMELINE
- Proposed in 2006 and approved in 2009
- Design and implementation: 2010-2013
- ICTS map: from 2014
- Digital Twin: 2024-2033

KPIs:
- >200 scientific publications, 2011-2023
- 20 European projects, 2014-2023
- 8 contracts with the private sector
- 8 collaborations with the public sector
- External financing > €6M
SOCIB: A multi-platform observing system in the western Mediterranean
SOCIB Data Quality Policy

- Follow **FAIR data principles**.
- Implement Quality Control, following international standards, both in real time and delayed time.
- Meet scientific and operational requirements of both the SOCIB multiplatform system and forecasting systems (including research projects in which SOCIB is involved).
- Include data assets in the **SOCIB Data Repository**.
- Ensure the compliance of SOCIB metadata with international standards.

Data Quality Strategy

01 Quality Assurance
02 Quality Control
03 Quality Assessment
04 Tools
05 DMPs
06 Data Profiling
07 Data Cleansing
08 Data Operations (SOPs)
09 Quality Control WG
SOCEB Data Management Framework

Raw/Processed RT data

Curated data

Archive, Open Access, FAIR

SOCEB DATA REPOSITORY
Data Management System: data products packaging

Data products and versioning system

Semantic versioning schema:
- $x.y.z = \text{major.minor.patches}$
- patches: new data
- minor: new variables
- major: (meta)data change

Operational data
- L0 netCDF
- L1 netCDF

Data product design: PI + DCF
- L2 netCDF
- L1 DM netCDF

Assessment and corrections

Publication

Curation: e.g. delayed mode

Packaging

DOI minting

Versioning

Data Product v1.0.0
### Operational data

**thredds.socib.es**

Data server: the structure and organization of the data responds to purely operational criteria. It also offers access through standard services (OPenDAP, WMS).

### Data Catalog

**apps.socib.es/data-catalog**

Catalog of data products: operational data in the best available quality, packaged according to scientific criteria. Data products have a DOI assigned to them.

### API M2M

**api.socib.es**

Machine-to-machine interface REST API to access repository data and metadata.
We promote **accessible** scientific **research** and **data** open to all citizens.

**FAIR**

**OPEN ACCESS**

**CORE TRUST SEAL**

**SOCIB Data Repository**

**TRUST**

**SOCIB Data Repository: support Open Science**
Data Center Facility: high level overview

- Observability
- QA
- Use
- Added value
- QC
- DATA MANAGEMENT SYSTEM
- Products Services
- Data & Information & Metadata
- Software Management
- IT & Software Architecture
- Monitoring
SOCIB Data Repository: link to EOSC

From local and regional to European and Global impact
Distinct from peer initiatives that focus on the human scholar, the FAIR Principles put specific emphasis on enhancing the ability of machines to automatically find and use the data, in addition to supporting its reuse by individuals (Wilkinson, M. D. et al., 2016).
Thanks a lot!