Advanced Digital Infrastructures and AI-Ready Workflows for Marine Research and Operational Services: From FAIR Data and HPC to the OneHealth AI Factory



Gonzalo González-Nuevo, Sofía González-Pérez, Adrián Sanjurjo-García, Luz M. García-García, Daniel Rodríguez-Abal and Manuel Ruiz-Villarreal

GOBIERNO DE ESPAÑA
MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES
CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS
INSTITUTO ESPAÑOL DE OCEANOGRAFÍA

INSTITUTO ESPAÑOL DE OCEANOGRAFÍA (IEO-CSIC)

¿QUÉ ES EL IEO?

CENTRO NACIONAL DE REFERENCIA EN CIENCIAS MARINAS (DESDE 1914) | INTEGRADO EN EL CSIC

- GENERADOR DE CONOCIMIENTO CIENTÍFICO
- ASESORAMIENTO CIENTÍFICO A ADMINISTRACIONES
- REPRESENTACIÓN INTERNACIONAL



ÁREAS DE INVESTIGACIÓN Y FUNCIONES

1 RECURSOS PESQUEROS SOTENIBLES

- Evaluación de caladeros
- Gestión sostenible
- Estudios biológicos pesqueros



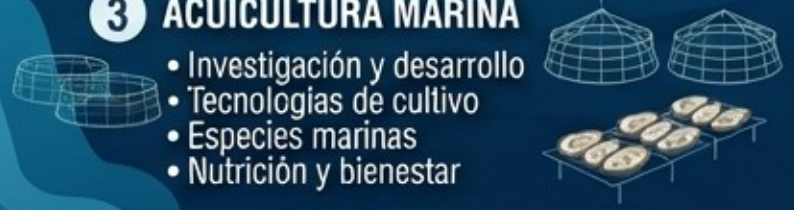


2 MEDIO MARINO Y PROTECCIÓN

- Cambio climático
- Oceanografía física, química y biológica
- Biodiversidad
- Contaminación

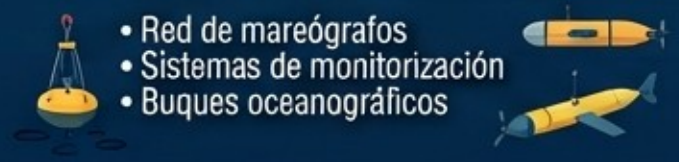
3 ACUICULTURA MARINA

- Investigación y desarrollo
- Tecnologías de cultivo
- Especies marinas
- Nutrición y bienestar



4 OBSERVACIÓN Y TECNOLOGÍA

- Red de mareógrafos
- Sistemas de monitorización
- Buques oceanográficos



CAPACIDADES OPERATIVAS

FLOTA OCEANOGRÁFICA



R/V RAMÓN MARGALEF



R/V ÁNGELES ALVARIÑO

Campañas en alta mar
Tecnología avanzada
Muestreo global

CENTROS COSTEROS



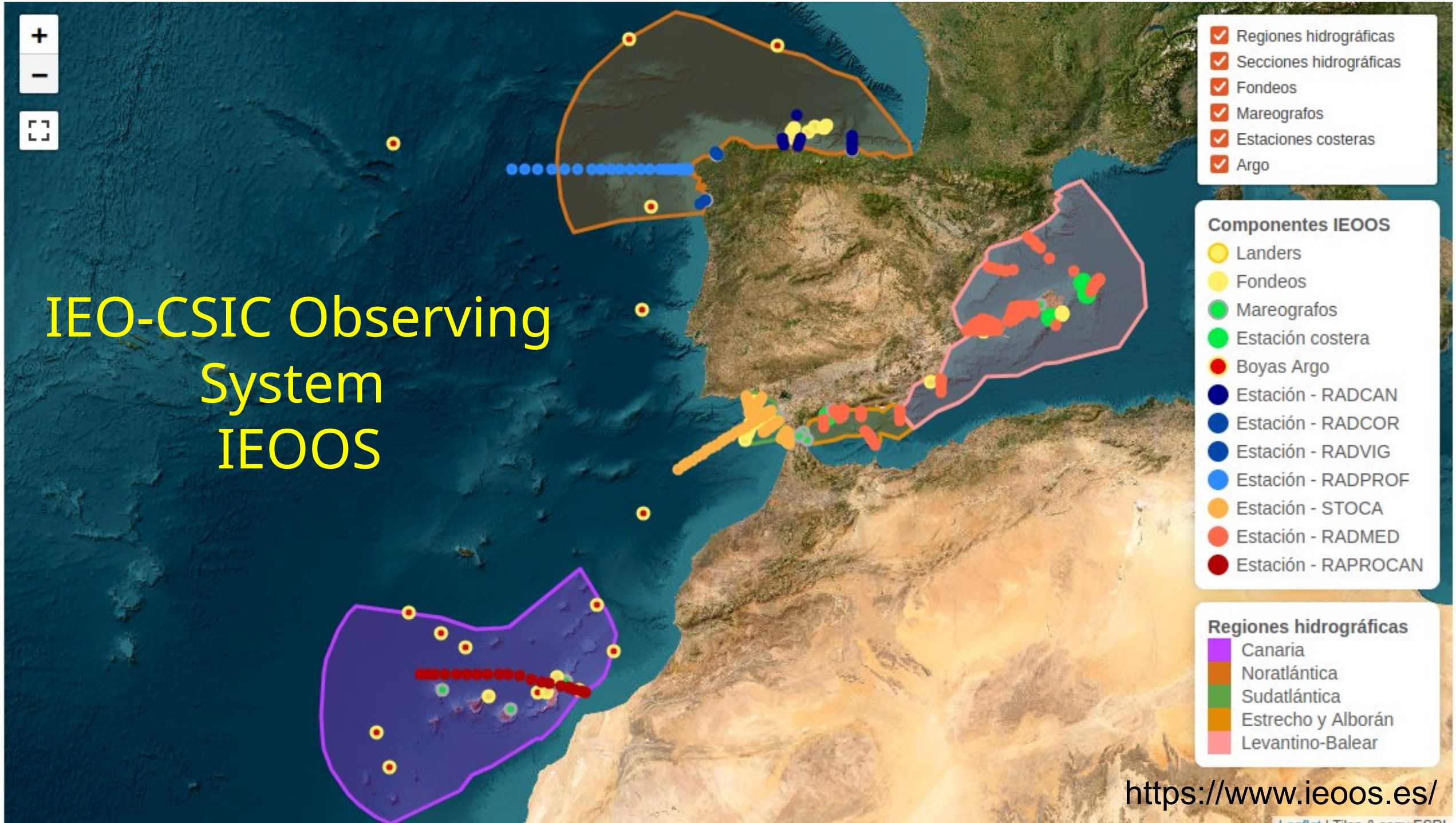
DIVULGACIÓN Y DATOS



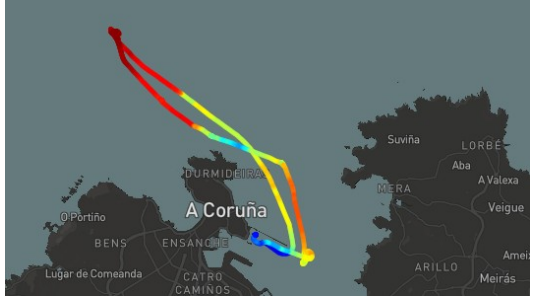
- Repositorio digital (e-IEO)
- Formación de investigadores
- Cultura marina

IMPACTO Y FUTURO

ECONOMÍA AZUL SOSTENIBILIDAD CONSERVACIÓN MARINA


IEO-CSIC Observing System IEOOS
Regiones hidrográficas
Secciones hidrográficas
Fondeos
Mareografos
Estaciones costeras
Argo
Componentes IEOOS
Landers
Fondeos
Mareografos
Estación costera
Boyas Argo
Estación - RADCAN
Estación - RADCOR
Estación - RADVIG
Estación - RADPROF
Estación - STOCA
Estación - RADMED
Estación - RAPROCAN
Regiones hidrográficas
Canaria
Noratlántica
Sudatlántica
Estrecho y Alborán
Levantino-Balear
https://www.ieoos.es/

Radiales





TSG



- Temperature
- Salinity
- Chorophyll A
- Oxygen
- CDOM
- Turbidity

Rosette-CTD



- CTD
  - Temperature
  - Salinity
  - Chlorophyll A
  - PAR
  - Oxygen
- Bottles
  - Nutrients
  - Chlorophyll
  - Oxygen
  - Ciliates
  - DNA
  - Citometry

Fito y Zooplankton



- Taxonomy
- Abundances
- Biomass
- FlowCAM
- Zooscan

PELACUS

Top Predators



Litter



Microplástics



CUFES



TSG



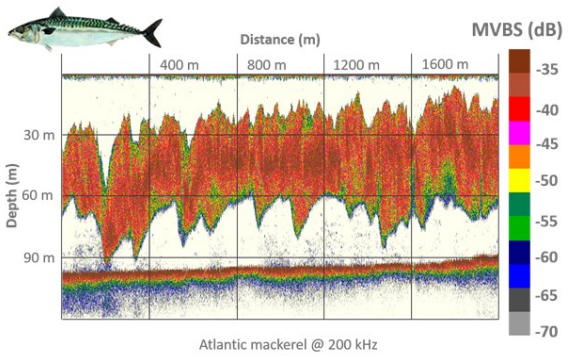
Meteorology



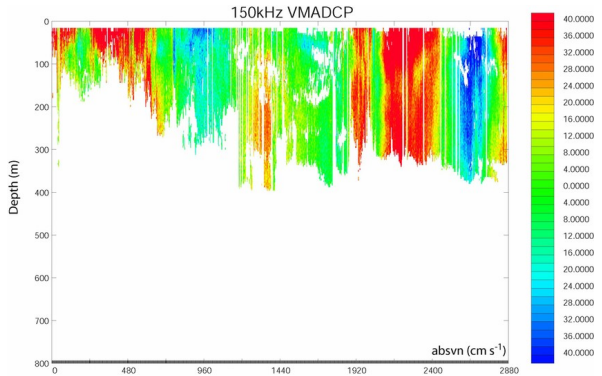
Fito y Zooplankton



Acustics



VADCP



Rosette-CTD



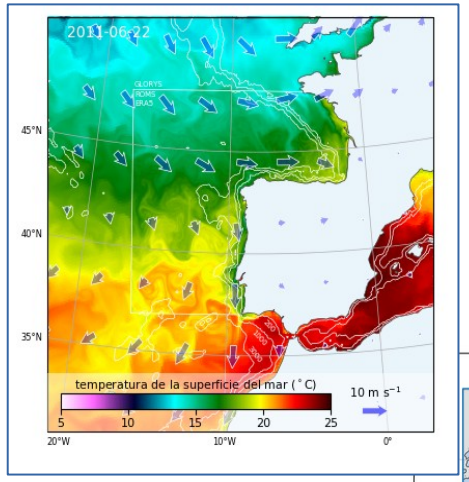
Hauls



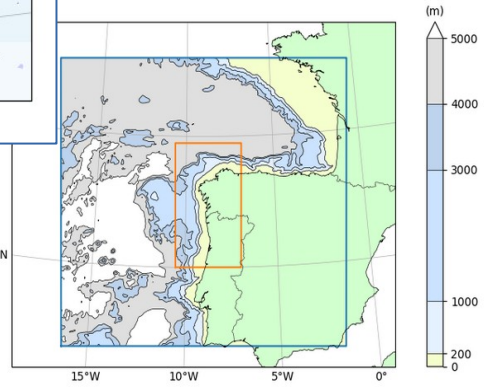
Ocean Models

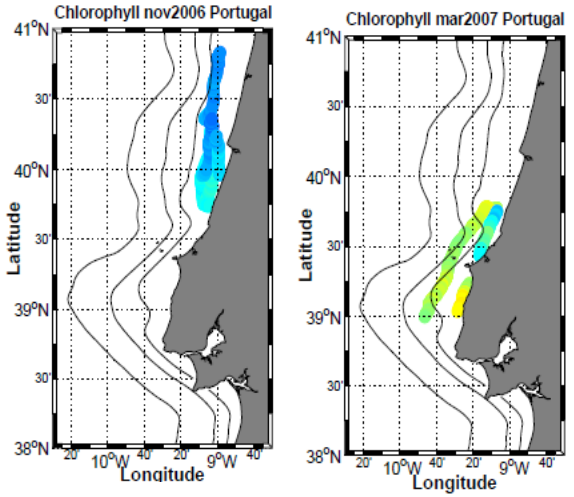
Hydrodynamic model (3D)
- ROMS Rutgers
- 30 vertical levels
- 3.5 km horizontal resolution
- Meteorological forcing
- Tides
- Rivers

Biogeochemical model N2PZD2 (nutrient sources)
- Fennel et al, 2006



Velocities
Turbulence
Temperature
Salinity
Density





Phyto
Zoo

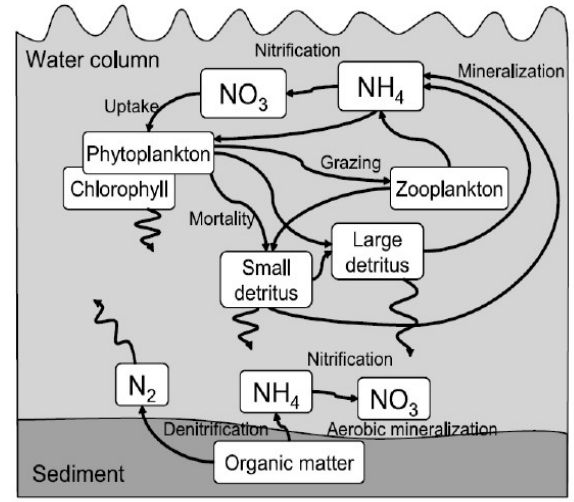
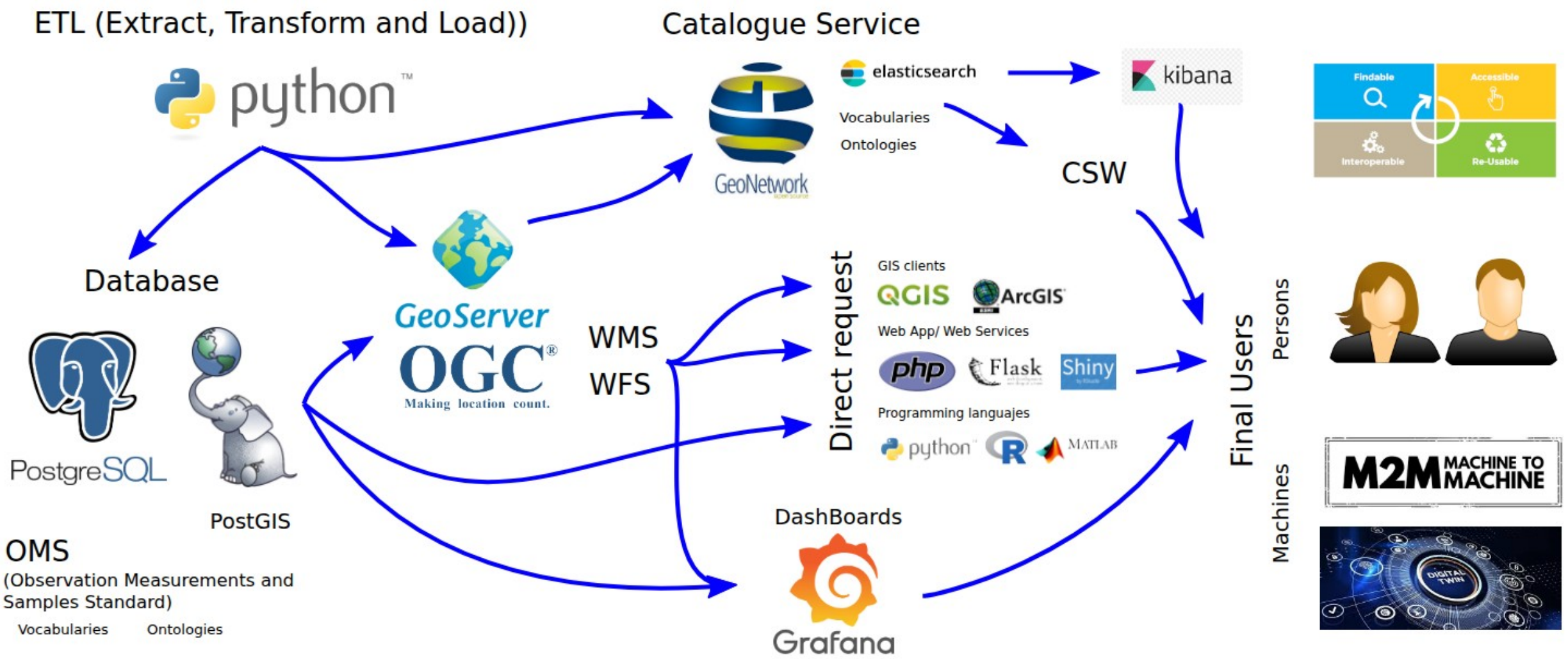


Figure 1. Biological model schematic.

Lagrangian IBM model
- Offline, Ichthyop or OpenDrift
- Physical processes (advection and dispersion)
- Biological behaviour of Early Life Stages (growth, buoyancy, vertical migration, etc.)

Transport
Survival
Connectivity

Data Ecosytem



FAIR data = Standards

- Findable
- Accesible
- Interoperable
- Reusable



Catalogue CWS



Node GIS WMS WCS WFS



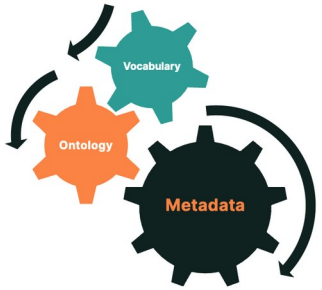
Files ISO19139 GeoJSON



Semantic Vocabularies

The only way to ensure that data is fair for machines—and ultimately for AI—is through the use of standards

Vocabularies and Ontologies



- Ontology: A formal definition of the types, properties, and relationships among entities that exist within a particular domain.
  - https://earthportal.eu/
- Vocabulary: A list of standard terms used in a specific field.
  - https://vocab.nerc.ac.uk/
  - https://vocab.seadatanet.org/search

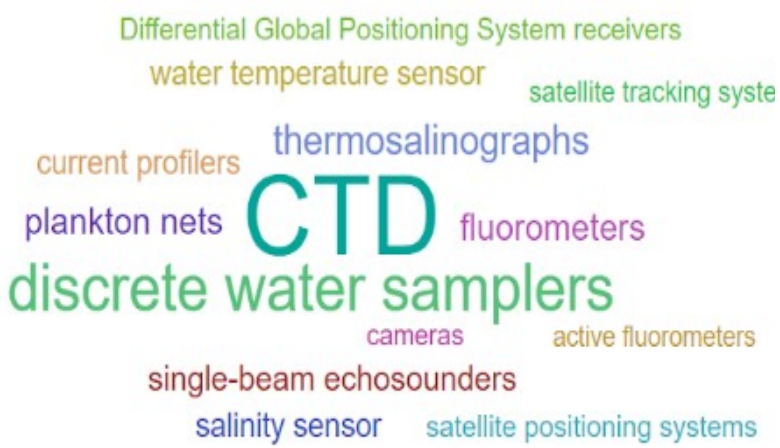
The other essential requirement for data to be used by machines without supervision is the use of standard vocabularies when creating metadata.

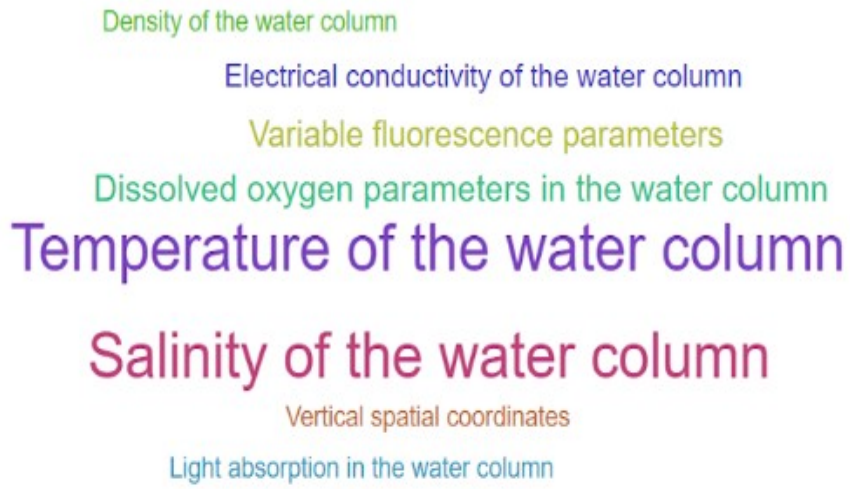
NODC IEO Geoportal

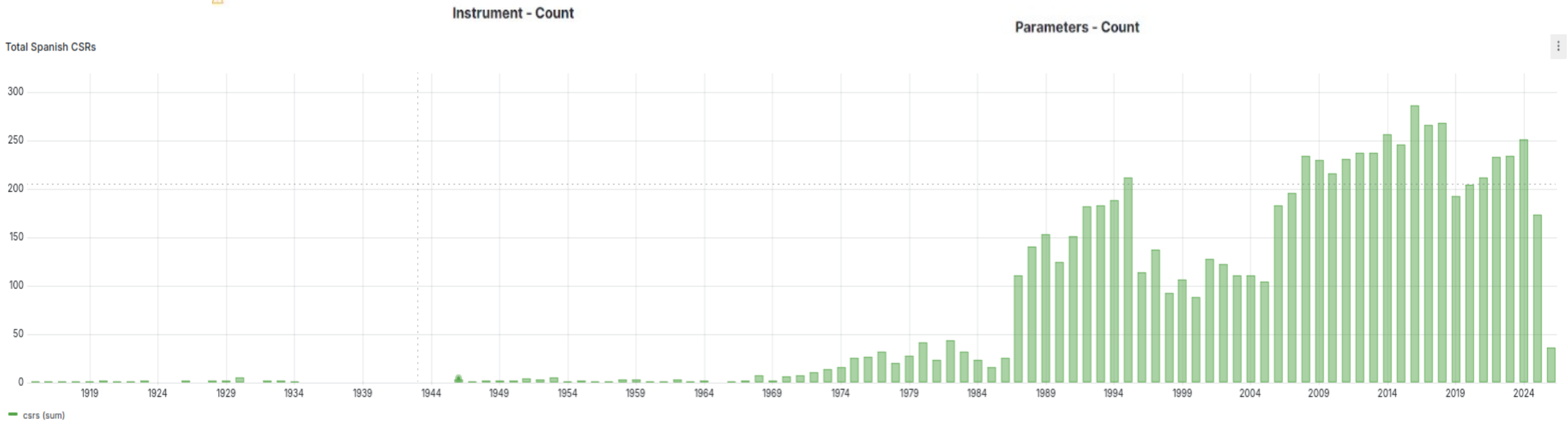


https://nodc.ieo.es/geonetwork

- It has a user-friendly interface
- Machines can query the cataloged information using standard CSW services or through its API.







CSIC
Consejo Superior de Investigaciones Científicas

INSTITUTO ESPAÑOL DE OCEANOGRAFÍA

SeaDataNet
PAN-EUROPEAN INFRASTRUCTURE FOR OCEAN & MARINE DATA MANAGEMENT

unesco
Intergovernmental Oceanographic Commission

IODE
International Oceanographic Data and Information Exchange

OBSERVATION AND MONITORING OF THE MARINE AND COASTAL ENVIRONMENT

- WP2 Marine governance and monitoring in Galicia
- WP3 New technologies and observation tools
- WP4 Integrated marine data platform
- WP5 Advanced analytical tools

CIENCIAS MARIÑAS GALICIA

SMART AND PRECISION SUSTAINABLE AQUACULTURE

SOCIAL AND ECONOMIC TRANSFORMATION FOR RESILIENCE

Xunta de Galicia

Sistema Universitario de Galicia

Consello Superior de Investigacións Científicas

202 RESEARCHERS

11 INSTITUTIONS

6 MILLION €

OCT 2022 - SEP 2025

CESGA
CETMAR
CIMA
CMAR - Xunta
Intecmar
MeteoGalicia

USC
UVigo
UDC

IIM - CSIC
IEO - CSIC

CIENCIAS MARIÑAS GALICIA

INTEGRATED MARINE OBSERVATION AND MONITORING STRATEGY, GALICIA

New observation technologies, tools and platforms

INTEGRATED MARINE DATA PLATFORM (Public, FAIR and connected)

Contributing from Galicia to the digital twin of the ocean



PCCMM x CESGA BigData



Marine Virtual Lab

Positioning Galicia as a leader in the management of marine data

Financiado por la Unión Europea
NextGenerationEU
PRTR PRTMPA
GOBIERNO DE ESPAÑA
MINISTERIO DE CIENCIA E INNOVACIÓN
Plan de Recuperación, Transformación y Resiliencia
XUNTA DE GALICIA

Integrated Data Platform for the Integration and Interoperability of Marine Data



- Do not collect all data
- Select different types of data
- V1 with more mature datasets
- V2 added images, videos, and DNA

Características

El Poder De Los Datos

Nuestra plataforma BigData permite a los usuarios desarrollar sus propios códigos de análisis y ejecutar tareas lanzando notebooks de Jupyter.

Acceso a datos científicos
Acceso a conjuntos de datos científicos marinos para su exploración y análisis.

Notebooks personalizados
No pierdas tu trabajo. Crea y guarda tus propios notebooks de Jupyter.

Infraestructura sólida
Utiliza una infraestructura robusta que garantiza un rendimiento óptimo y fluido.

Análisis avanzado de datos
Realiza análisis profundos de los conjuntos de datos científicos disponibles.

CIENCIAS MARIÑAS GALICIA · CESGA GALICIA SUPERCOMPUTING CENTER

Arquitectura da plataforma BigData

Autores
Pablo Prieto Rúa, Javier Cacheiro López, Cecilia Grela Llerena, Carlos Fernández Sánchez, Francisco Landeira Vega



Technology Challenge

#CiênciasMariñasGal

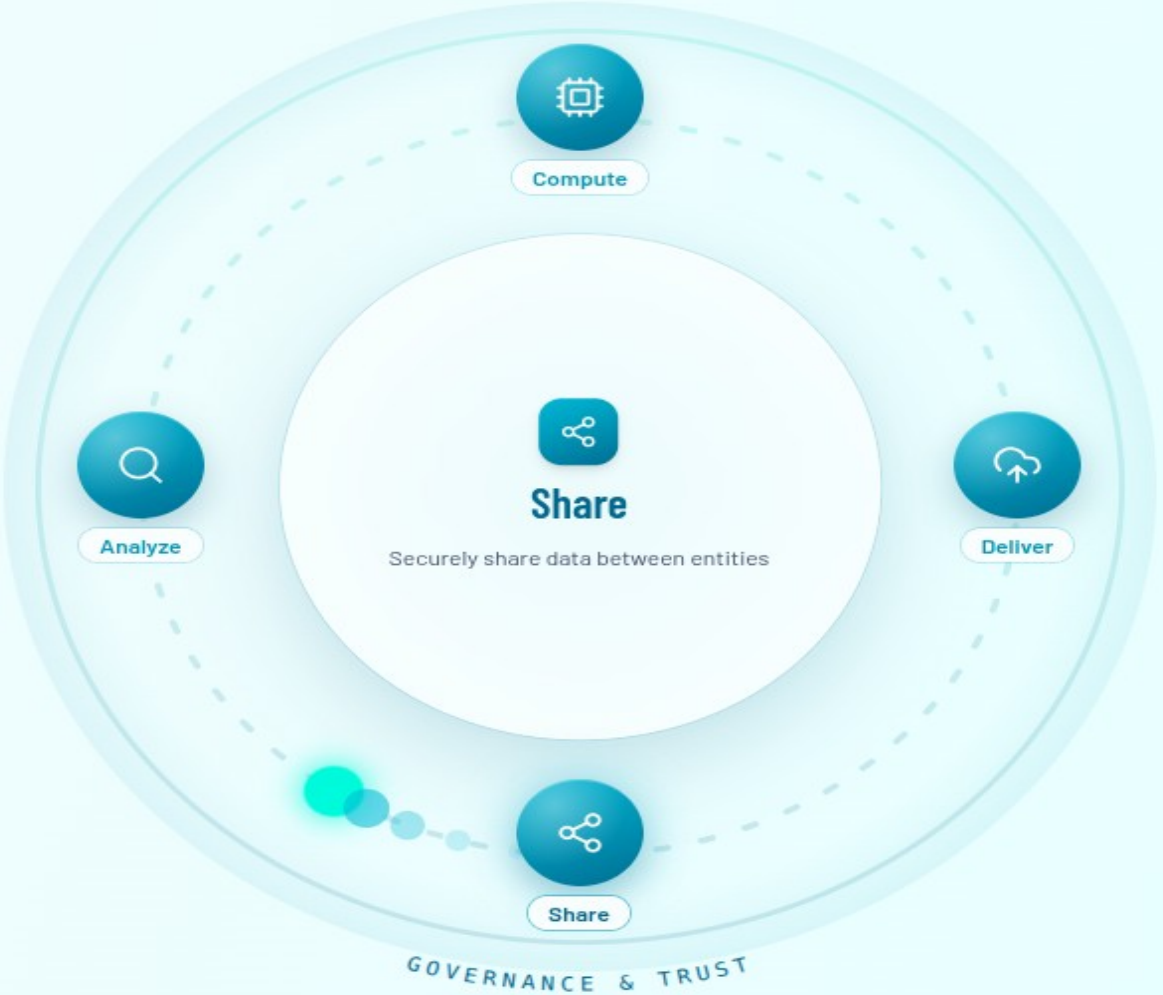
Financiado por la Unión Europea
NextGenerationEU
FEMP-FEMPA

Gobierno de España · Ministerio de Ciencia, Innovación y Universidades · Plan de Recuperación, Transformación y Resiliencia · Xunta de Galicia

More than data

OneHealth DataSpace goes beyond data sharing.
Analyze, process, and publish your data on the same platform, with full control throughout its lifecycle.

Access the platform ↗ Learn more



Risk of Portuguese man-of-war on Galician beaches


EL PAÍS

1 año, 12 € ¡Suscríbete!

Clima y Medio Ambiente

CAMBIO CLIMÁTICO · MEDIO AMBIENTE · EL TIEMPO · ÚLTIMAS

MEDUSAS >

Las peligrosas carabelas portuguesas echan el ancla en las playas españolas

Hasta ahora se veían unas semanas a principios del verano, pero la presencia a estas alturas de estos animales, que causan graves picaduras, desconcierta a los científicos, que apuntan al cambio climático


EL BOTIQUÍN

Carabelas portuguesas en Galicia: qué son, últimos casos y cómo actuar si te pican

LA VOZ DE LA SALUD

Carabela portuguesa flotando en el mar cercano de las Illas Cíes Manuel E. Garci UMO - IIM - CSIC

SOMOSMAR

Las carabelas portuguesas llegan a las islas Cíes y a Ons

LA VOZ VIGO


Lucio

Cientos de turistas siguen yendo al parque nacional. Los expertos recomiendan no tocarlas por su alta toxicidad

07 sep 2025 . Actualizado a las 13:41 h.


GALICIA

Las carabelas portuguesas llegan a diferentes playas de A Coruña y Carballo

Por ahora, no se ha prohibido el baño en las playas coruñesas

2 comentarios

- Physalia physalis (commonly known as the “Portuguese man-of-war”) appeared on the beaches of Galicia in late summer 2025.

- Physalia physalis causes a very painful toxic sting.

- Our app aims to predict its arrival on the Galician coast so that end users can stay informed about these phenomena.

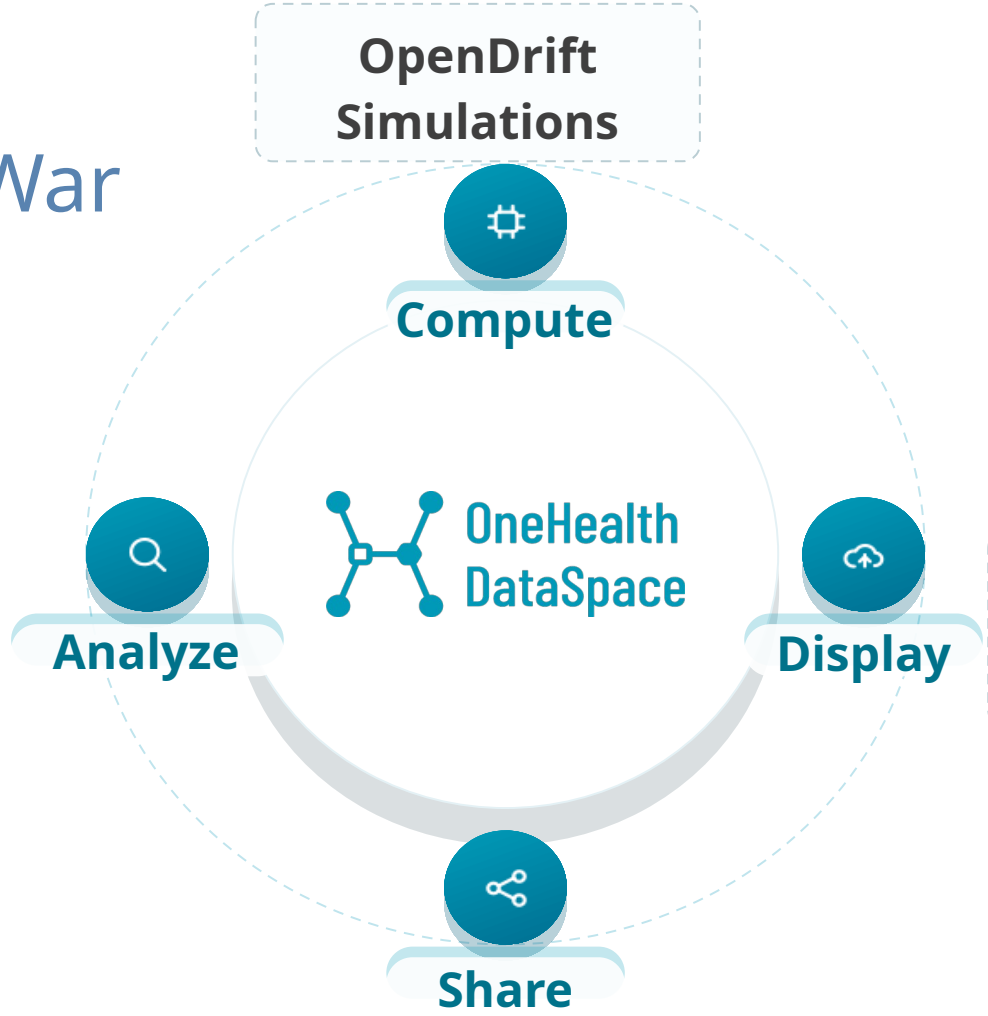
- Our app combines numerical simulations with observational data to provide comprehensive and easy-to-understand information.

- Our app is designed to be easy to use for any user.

Estimating the Probability of Portuguese Man-of-War Sightings on the Beaches of Galicia

INSTITUTO ESPAÑOL DE OCEANOGRAFÍA

Data Analysis ROMS (MeteoGalicia)
Data Analysis Radar HF (Intecmar)



BD: PostgresSQL
Monitorization: Graphana
Web: carabelas.dataspace.cesga.es

Assets used: ROMS (MeteoGalicia) y Radar HF (Intecmar)
Assets produced: OpenDrift Trajectories and Physalia Physalis Risk Assessment

carabelas.dataspace.cesga.es

Importar marcadores… | Comenzar a usar Firefox | Máis visitados | Getting Started | Comenzar a usar Firefox | Inicio - MYCOAST | Iniciarse | Frontiers | Current Sta… | PROYECTOS DE GENE… | choi | Outros marcadores

Carabelas · Praias de Galicia

Que é? GL ES EN

Buscar praia...

Praias preto de min

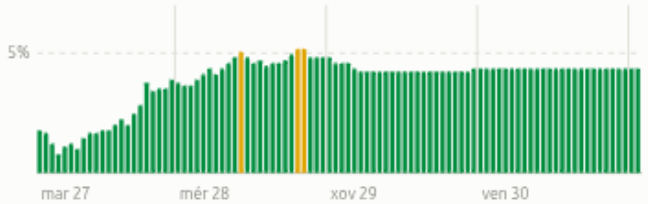
Beo

Nivel máximo previsto (4 días)

Moderado 5.2%

Mar 27	Mér 28	Xov 29	Ven 30	Sáb 31
3.9%	5.2%	4.8%	4.4%	4.4%
Baixo	Moderado	Baixo	Baixo	Baixo

Evolución nos vindeiros 4 días



Máximo horario sobre todos os modelos e isóbatas (peor caso)

▸ Detalle por modelo e isóbata

Recomendacións

- Extrema a precaución no baño e vixía a superficie da auga.
- Non toques exemplares varados e mantén a distancia coas crianzas e mascotas.

☎ En caso de picadura grave chama ao 112

Actualizado: 27 out, 01:00

Previsión baseada nun modelo probabilístico de deriva (peor caso). Non é unha garantía: segue sempre as indicacións do socorrismo.

Capa de partículas

Leeway | OceanDrift

Predición (+4 días) | Reanálise (−4 días)

▶ 22 xul, 16:00

Á deriva · Varada

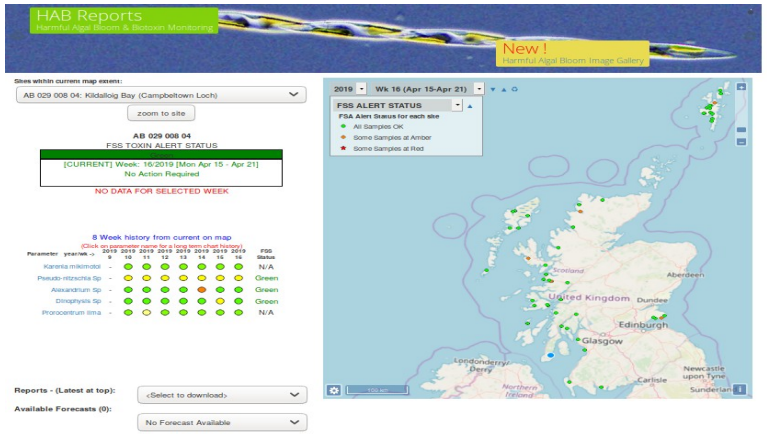
Nivel de aviso Sen risco · Baixo · Moderado · Alto · Moi alto

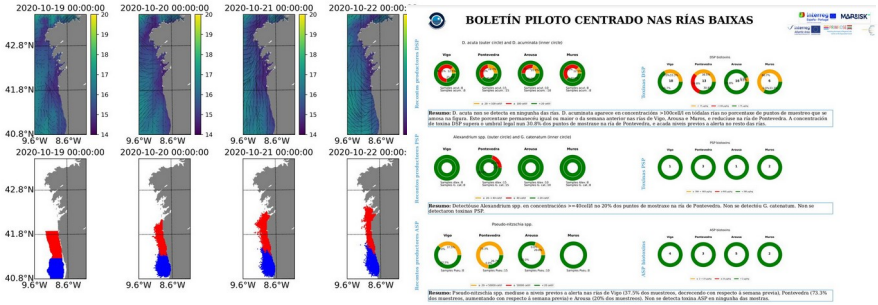
Actualizado: 21 xul, 02:00

Ocultar simulación

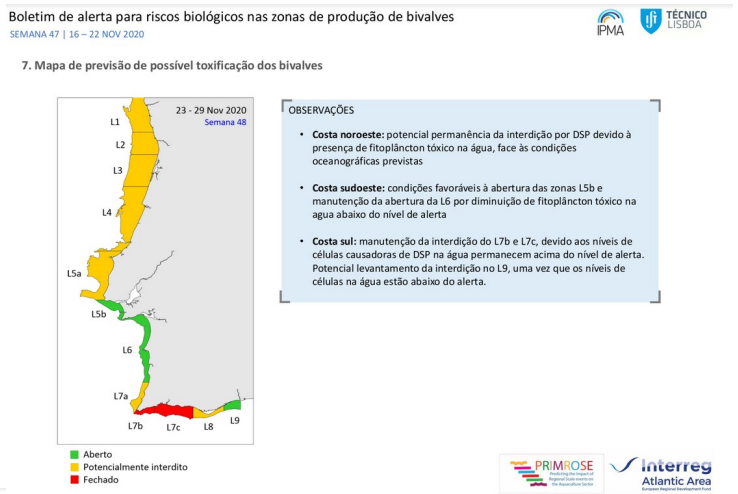
Cee · Noia · O Porto do Son · Boiro · Rianx · A Pobra do Caramiñal · Vilanova de Ar · Camba · O Grove · PONTEVEDRA

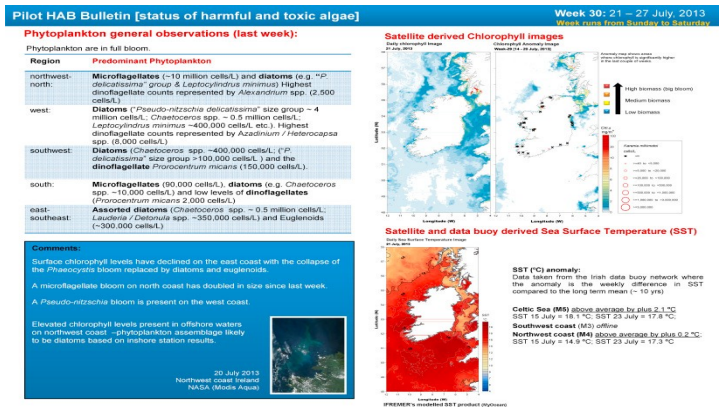
Early Warning System for Harnfull Algae Blooms

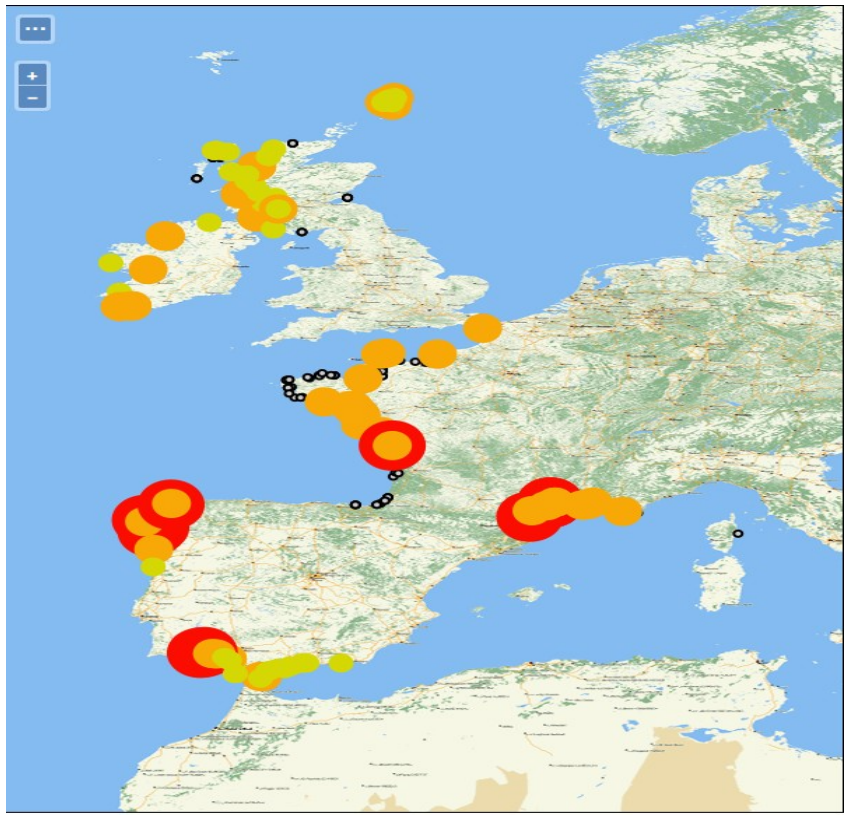












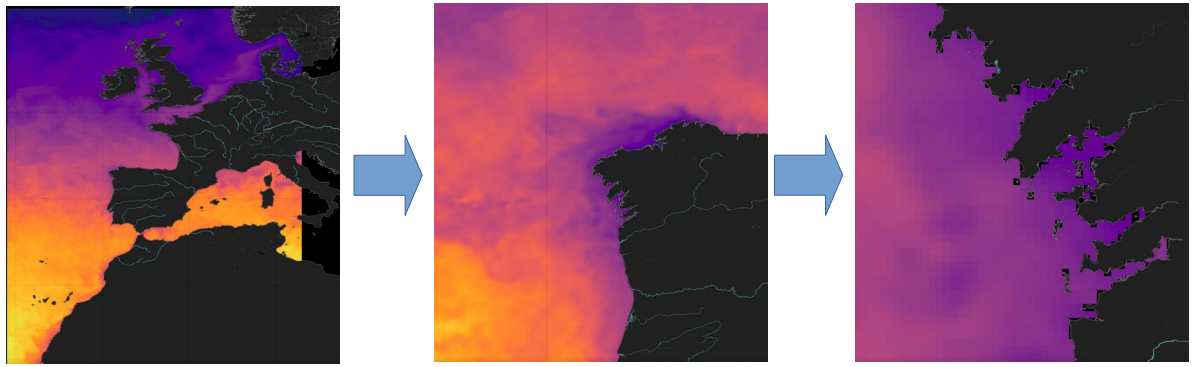
Dinophysis acuminata September 2020

Data sharing, Standards and interoperability

HAB early warning http://www.shellfish-safety.eu/

Next steps One Health AI Factory

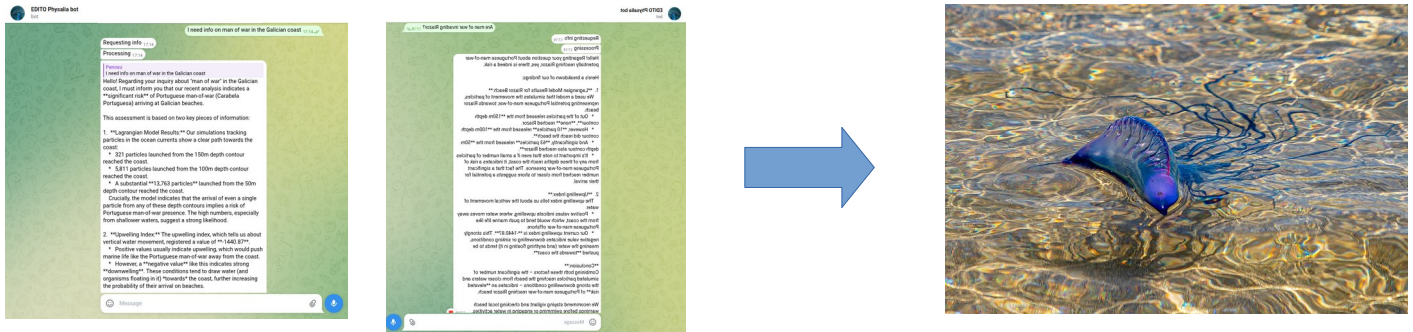
- Apply AI to downscale ocean circulation models in order to simplify the forcing of Lagrangian models.



- Extract unstructured information from Portuguese man-of-war sightings to improve risk estimates.



- Use LLMs to improve the interaction between end users and the information generated by predictive systems.




N42.625°
N42.575°
25°
W8.875°
W8.825°
N42.525°
Thank you
for your attention
Data SIO, NOAA, U.S. Navy, NGA, GEBCO
Image © 2013 DigitalGlobe
Google earth
lat. 42.621656° long. -8.947970° elev. 95 m alt. ojo 5.59 km

FAIR data = Standards

- Findable
- Accesible
- Interoperable
- Reusable





Catalogue

CWS





Node GIS

STAC api
OGC api



Database,
Files, etc

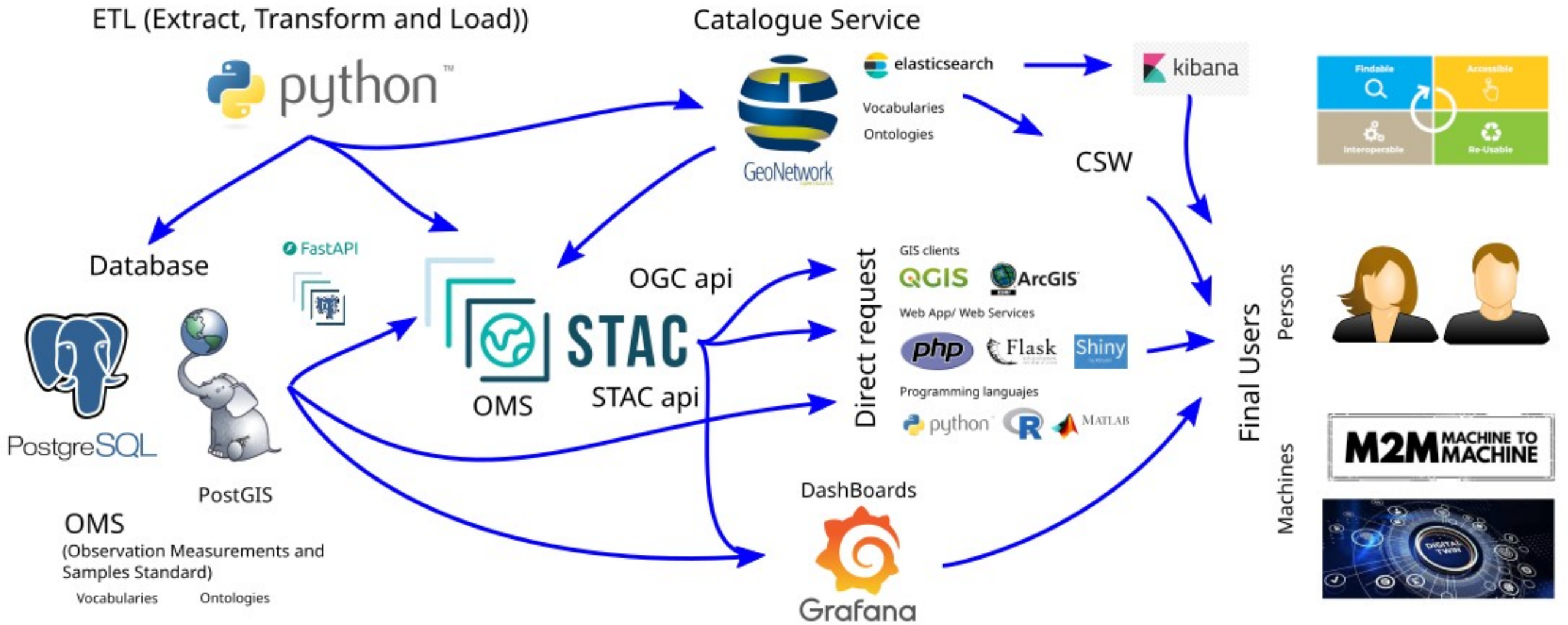
OMS, ISO19139
GeoJSON



Semantic

Vocabularies

Data Ecosytem + STAC



FAIR data = Standards

- Findable
- Accesible
- Interoperable
- Reusable





Catalogue

CWS







Cloud HPC

STAC api
OGC api



Database,
Files, etc

OMS, ISO19139
GeoJSON



Semantic

Vocabularies

Data Ecosytem + HPC

