



DATA
TERRA



ODATIS

Données et Services du CDS-IS-Coriolis

Thierry.Carval@ifremer.fr

Assemblée Générale ODATIS

12-13 décembre 2023

Observatoire Aquitain des Sciences de l'Univers, Pessac



Faits marquants

- Copernicus release 202311
- GDAC bouées dérivantes livraison à C3S pour ERA6
- Candidature GDAC bouées fixes et Argo
- En route vers OneArgo
- Bigdata Coriolis
- Projets européens
- Données Coriolis 2023
- Perspectives R&D format ARCO et IA
- Point budget, personnel
- Attentes du CDS-Coriolis vis à vis d'Odatis



Faits marquants 2023, Copernicus marine service

Copernicus Marine November 2023 release

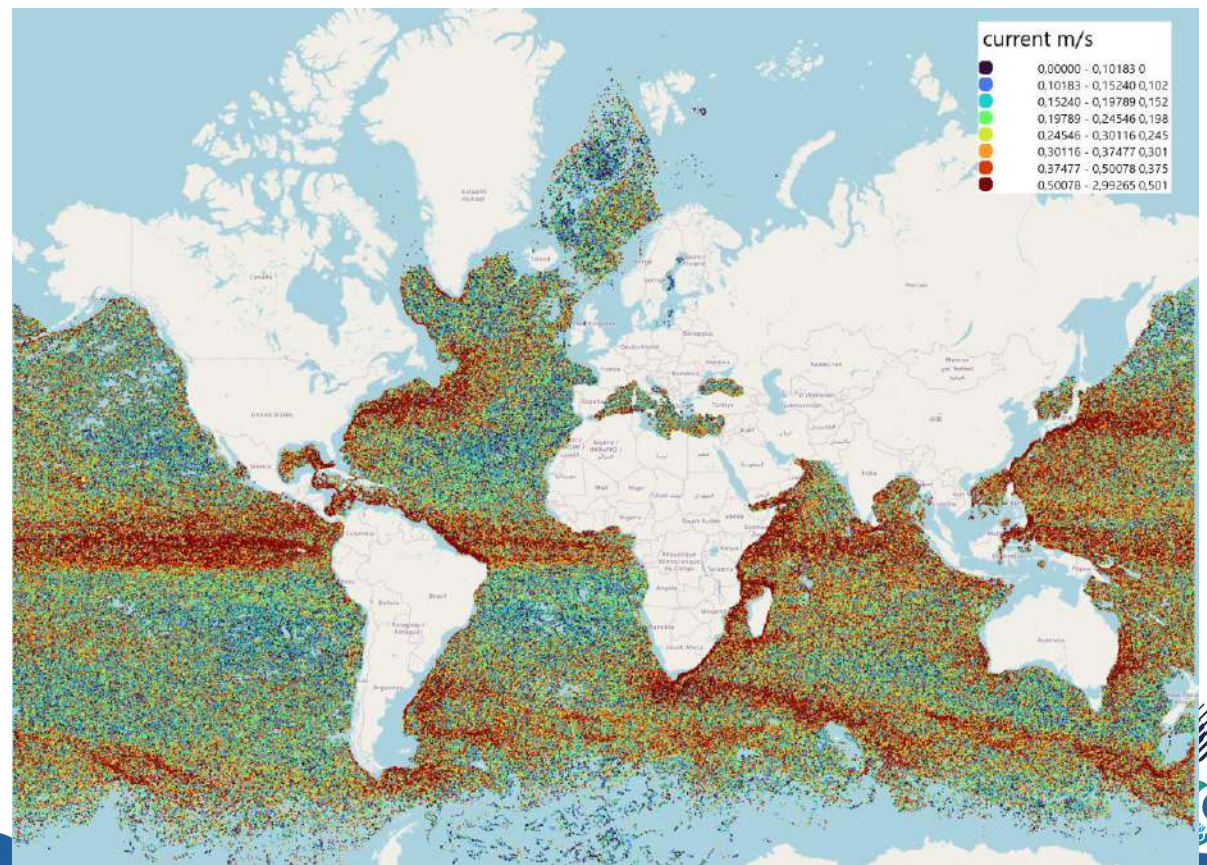
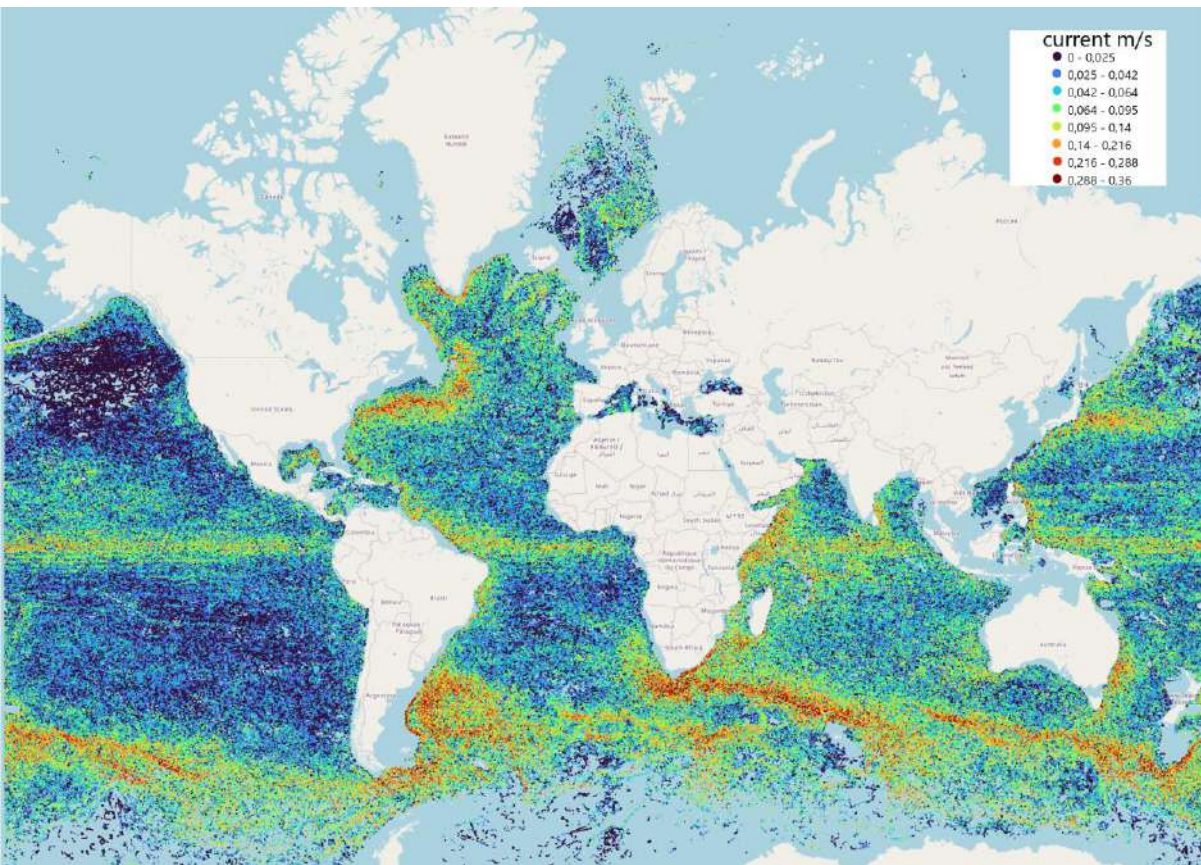
- Révision majeure du format NetCDF Copernicus Marine in situ V2
Révision des géométries, conformité CF 1.11 et ACDD
<https://doi.org/10.13155/59938>
- Temps différé retraitement global
 - T&S CORA (inclus correction Argo ASD avec un GHC cohérent)
 - Courants (bouées, adcp, radar hf, argo)
 - Vagues (bouées ancrées, dérivantes)
 - BGC (oxygène, chlorophylle)
 - Marégraphes
- Démarrage Copernicus MDS – Marine Data Store : cloud et bigdata
 - API données
 - Data lake Arco (zarr, parquet), S3, K8s



Faits marquants 2023, Copernicus marine service

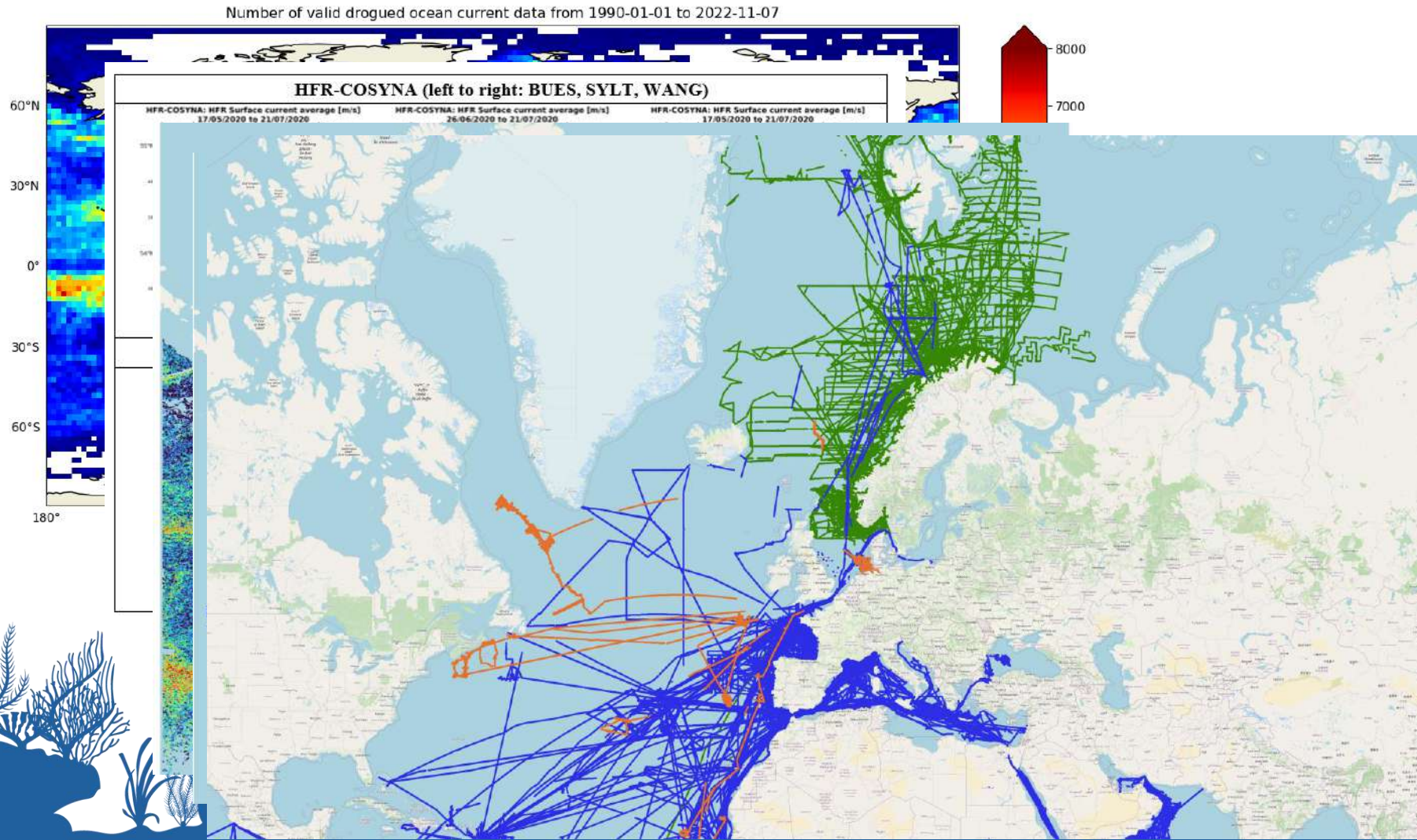
Copernicus Marine 202311 courants Argo, deux évolutions majeures temps réel

- Une série de 20 tests additionnels est appliquée à chaque trajectoire de flotteur Argo
Herbert Gaelle (2020). Qualification temps réel des données trajectoire des flotteurs Argo. <https://doi.org/10.13155/95169>
- Les courants TR sont calculés avec Andro (comme TD)
Ollivraut Michel, Rannou Jean-Philippe. ANDRO Dataset contents and format. <https://archimer.ifremer.fr/doc/00360/47126>



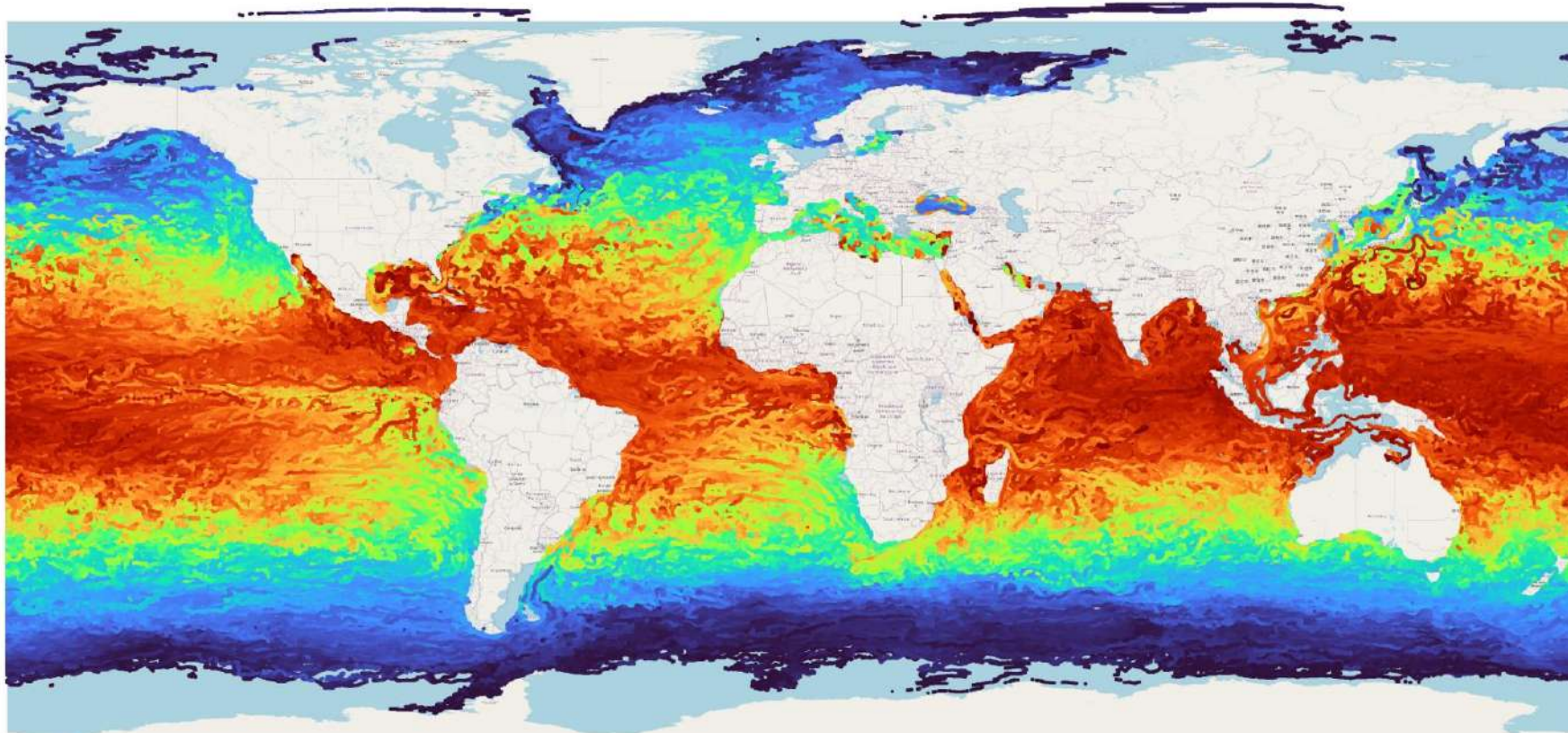
Faits marquants 2023 Copernicus marine service

- Copernicus Marine November 2023 delivery
 - Courants temps différé ADCP, Argo, radar HF, bouées dérivantes



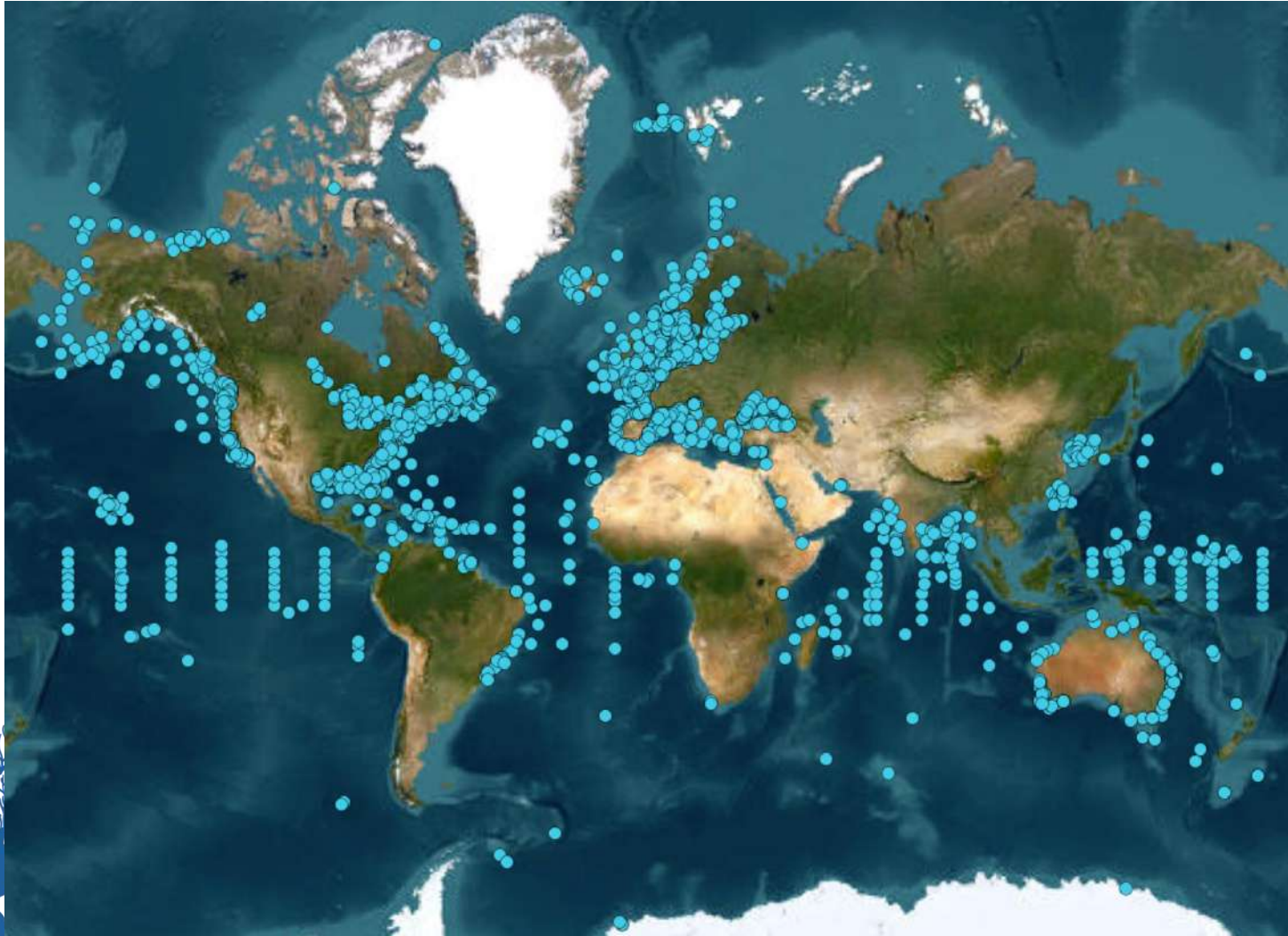
Faits marquants Copernicus COINS et C3S

- Copernicus COINS C-RAID : mise à jour en novembre 2023 (+2000 bouées)
- Copernicus Climate Change Service – C3S
Livraison du jeu de données de bouées dérivantes Coriolis au format BUFR
27 000 bouées, 12,6 millions de fichiers (un par jour)
Destiné à la réanalyse ERA6 (sst, sss, pression atmosphérique, vent, vagues)



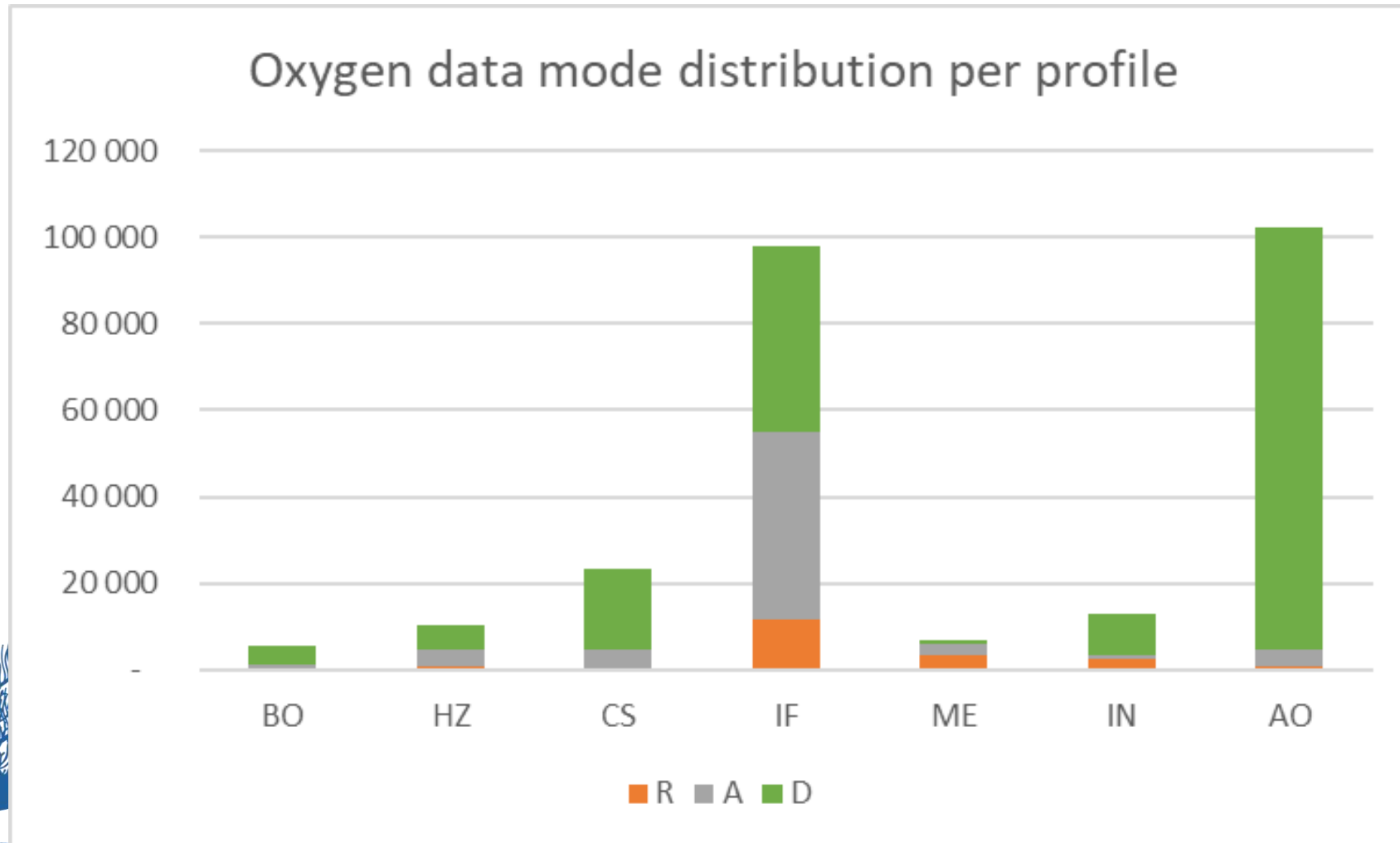
Faits marquants WMO-IOC

- Candidature officielle GDAC bouées fixes
- Candidature GDAC Argo à l'étude



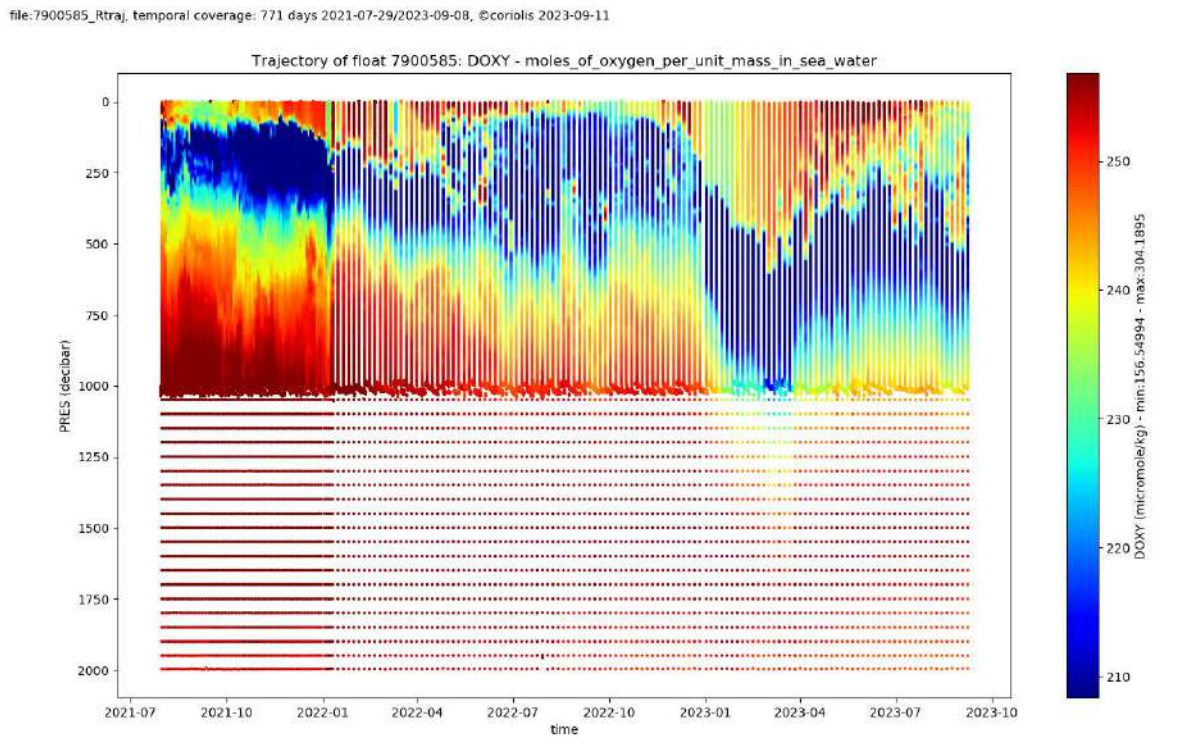
Faits marquants Argo

- Ticket en cours “doxy ajustement temps réel après 5 cycles »
- Un contrat pour temps différé oxygène des flotteurs Coriolis



Faits marquants Argo

- BBP: retraitement complet “particulate backscattering” avec la nouvelle procedure qualité <https://doi.org/10.13155/60262>
- CHLA: retraitement complet chlorophyle avec la nouvelle procedure qualité <https://dx.doi.org/10.13155/35385>
- Trajectoires: retraitement complet des 3700 flotteurs distribuées dans le nouveau riche format trajectoires V3.2 qui contient toutes les évènements datés du flotteur



Faits marquants Argo

- Nouvelle fiche de déploiement flotteurs Coriolis, améliore la saisie des métadonnées et calibrations (dont la richesse va croissant) avec champs et contrôles prédéfinis, partagée sur <https://cloud.ifremer.fr>
- Web d'aide à la récupération de flotteurs Coriolis régulièrement amélioré
A l'étude : diffuser un flux VMS des flotteurs à récupérer
<https://floatrecovery.euro-argo.eu/>



Faits marquants Argo

- La chaîne de traitement flotteurs Coriolis est distribuée sur GitHub Euro-Argo, 500 000 lignes de code
- The Coriolis BGC-Argo floats data processing chain is continuously improved (a monthly release), we now manage 84 versions within 5 families (Apex, Navis, Nemo, Nova, Provor).
- FAIR data : Coriolis Argo data processing chain is now on GitHub

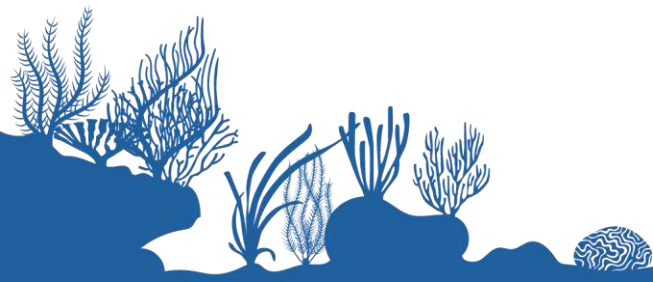
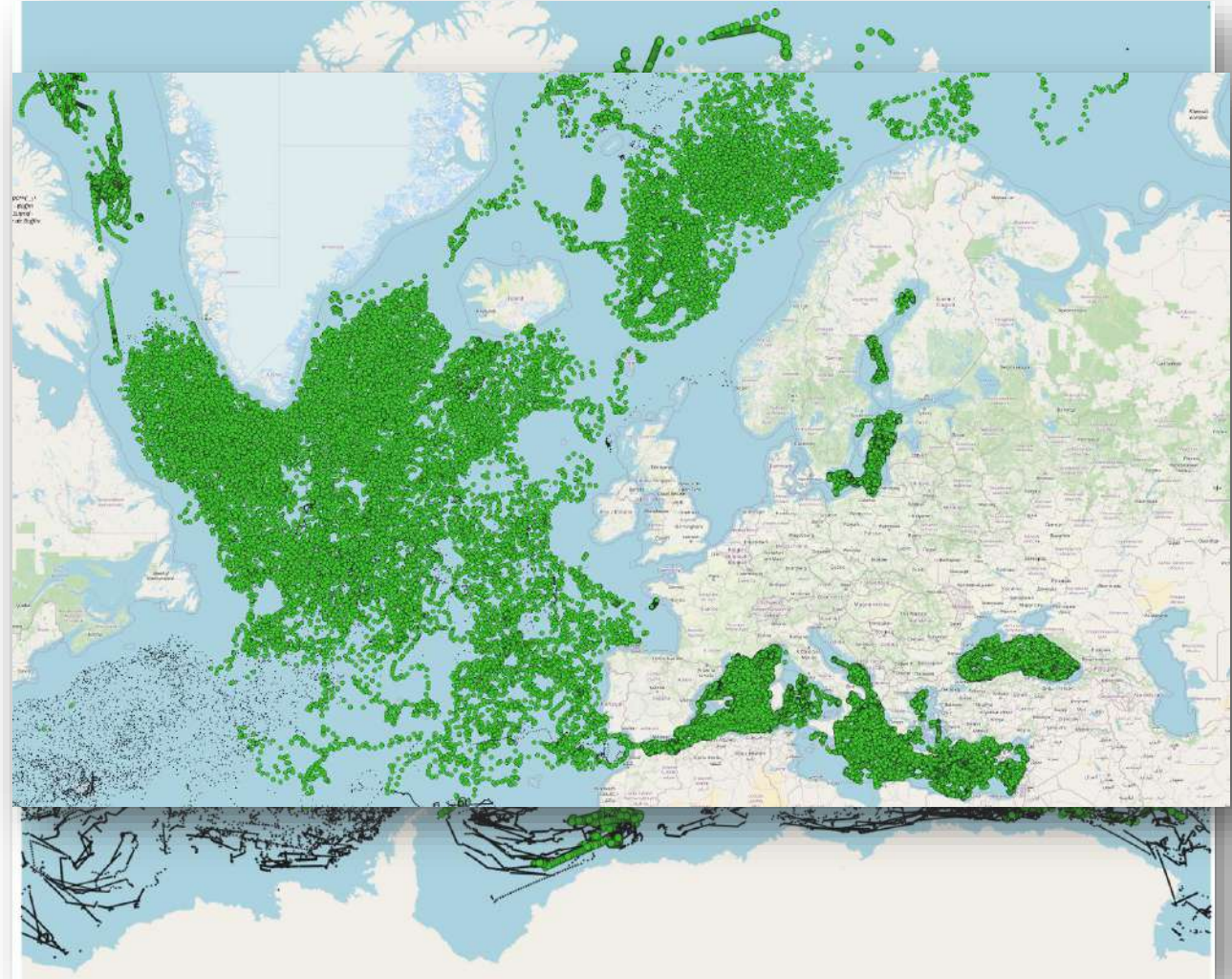


- <https://github.com/euroargodev/Coriolis-data-processing-chain-for-Argo-floats>
- 52 major releases, 323 tags (releases + patches)



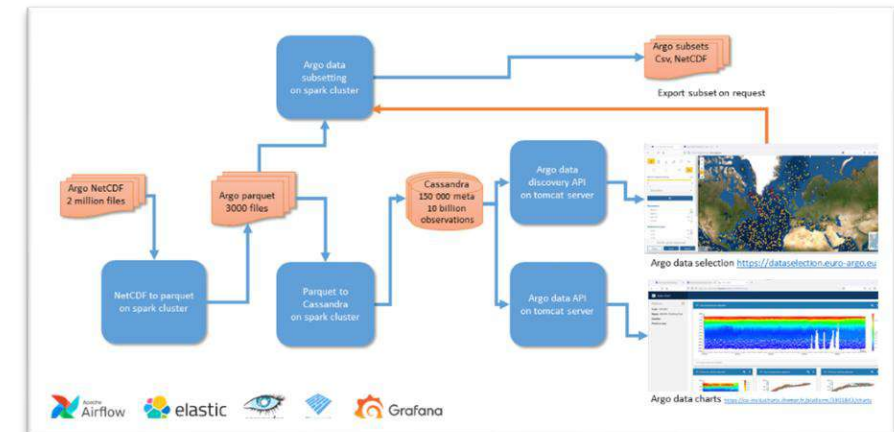
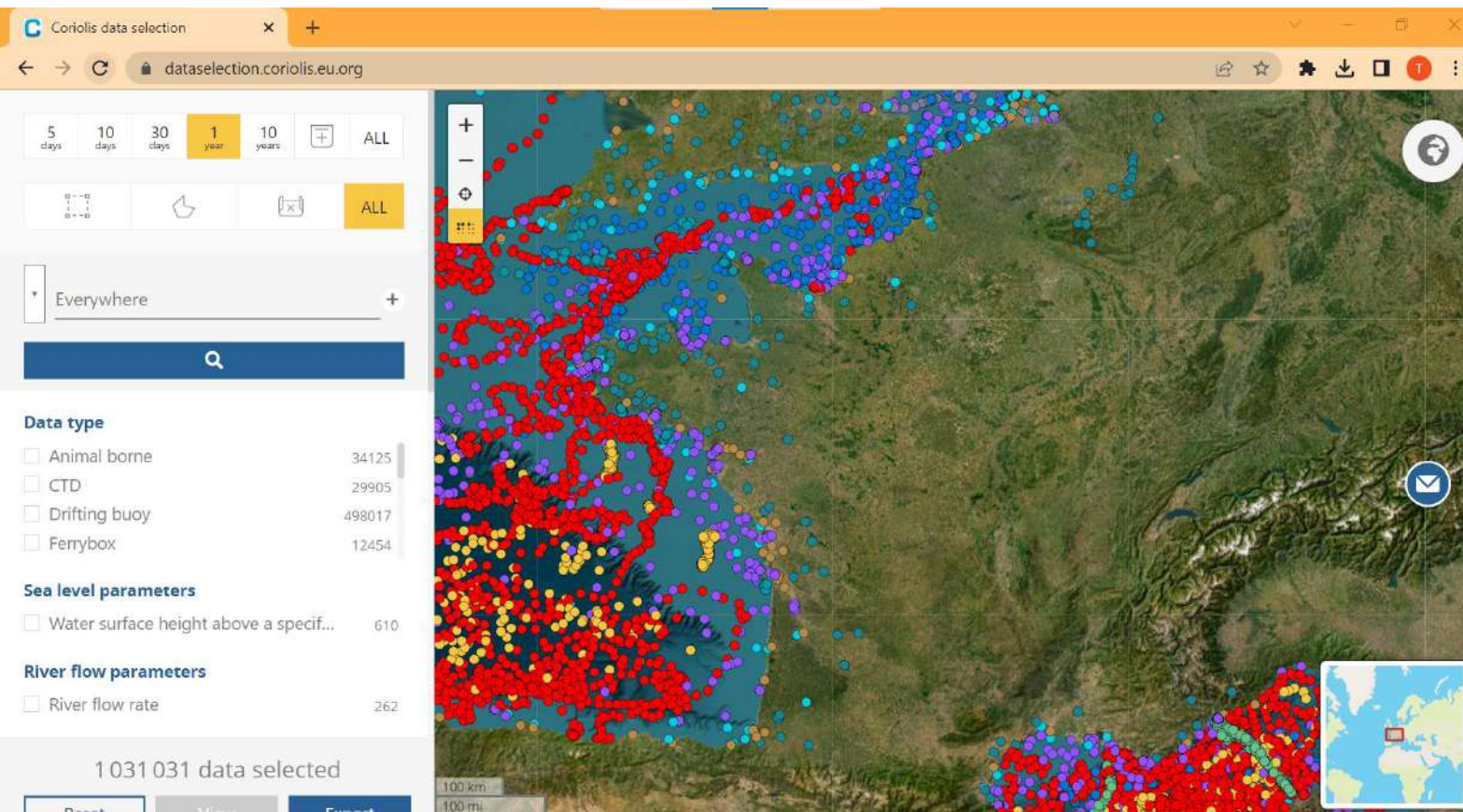
Faits marquants Argo

- The 733 Coriolis BGC-Argo floats map (green dots)
- A major data source for oxygen, chlorophyll, nitrate, pH, turbidity, optic properties, pCO₂
- Significant coverage of marginal seas
med sea, black sea, baltic, red sea, baffin sea



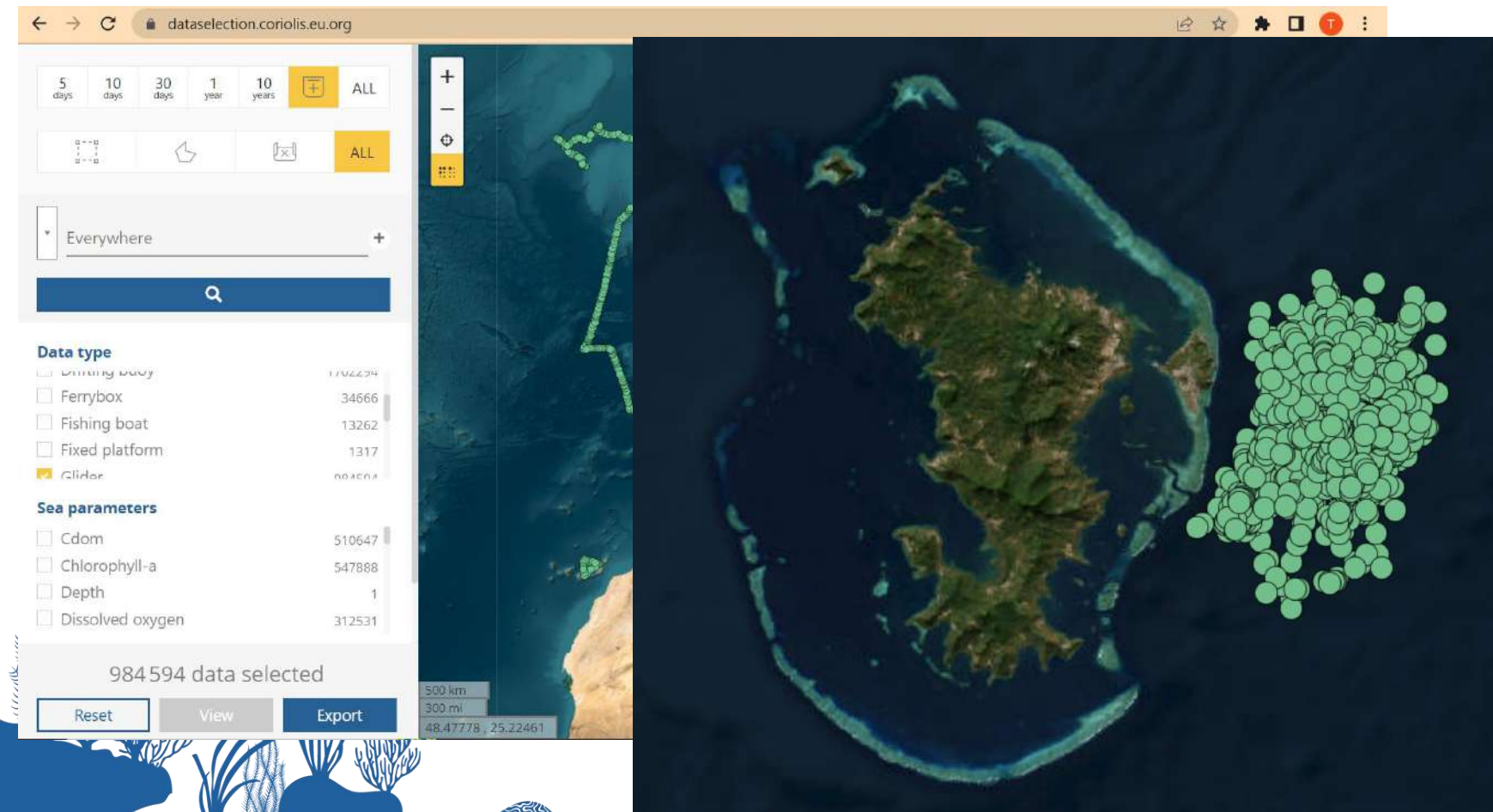
Faits marquants

- Découverte et subsetting données Coriolis, périmètres complet (profils verticaux, trajectoires et series temporelles)
- Prochaines étapes : le streaming (mise à jour au ¼ d'heure), les radars HF, les spectres de houle (bigdata spark-cassandra sur 20 million de fichiers)

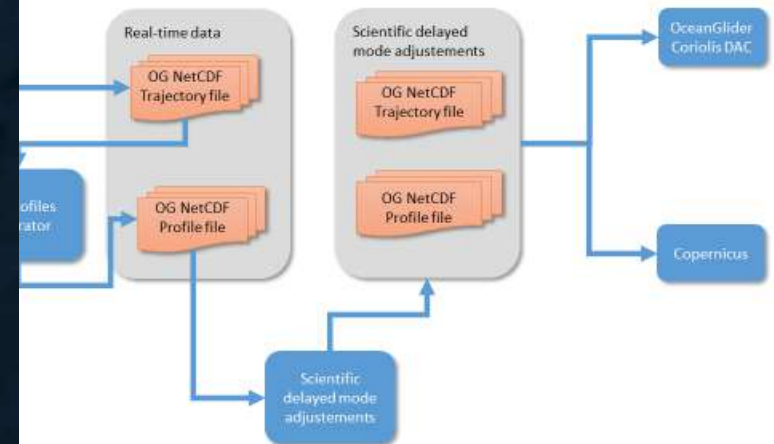


Faits marquants Gliders

- Spécification du workflow temps différé français
- Observation continue sur volcan de Mayotte depuis 2 ans, record d'endurance avec 44 000 profils verticaux (Seaexplorer)



flow



Faits marquants 2023 projets européens

Projets en cours

- EOSC-future : exploration interactive données Argo et SeaDataNet <https://eosc-future.maris.nl>
- EOSC Blue-Cloud : webinaire FAIR-DATA ERICs marins (Argo, EMSO, ICOS, SeaDataNet)
- EOSC Fair-Ease workbench BGC: Argo S3, ODV, Scoop-Argo, Diva, Jupyter notebooks sur Galaxy Earth <https://earth-system.usegalaxy.eu>
- EOSC Fair-Impact vocabulaires et ontologie Argo dans Onto-Portal <https://ontoportals.org/>
- EOSC UE George : QC oxygène, incertitudes Argo et ICOS-Global-Socat
- EuroGO-SHIP: organiser les flux de données Européens GO-SHIP (PI - NODC – SeaDataNet – OceanOPS - CCHDO - WOD - Glodap – Copernicus – GTS)
- Copernicus-MDS expertise in situ sur le cloud Copernicus
- Copernicus COINS temps différé bouées dérivantes, stratégies d'observations
- [Copernicus in situ – Trusted](#) : bouées de références pour SST (Sentinel, GHRSSST)
- Argo GDAC sur ASDI, le cloud Amazon environnement, diffusion quotidienne <https://registry.opendata.aws/argo-gdac-marinedata>



Faits marquants 2023 projets européens

Projets acceptés

- AMRIT : renforcer l'activité DAC-GDAC des ERICs (activité amont)
- ENVRI-Hub-Next interopérabilité des ERICs (activité aval)

Projets en cours de montage

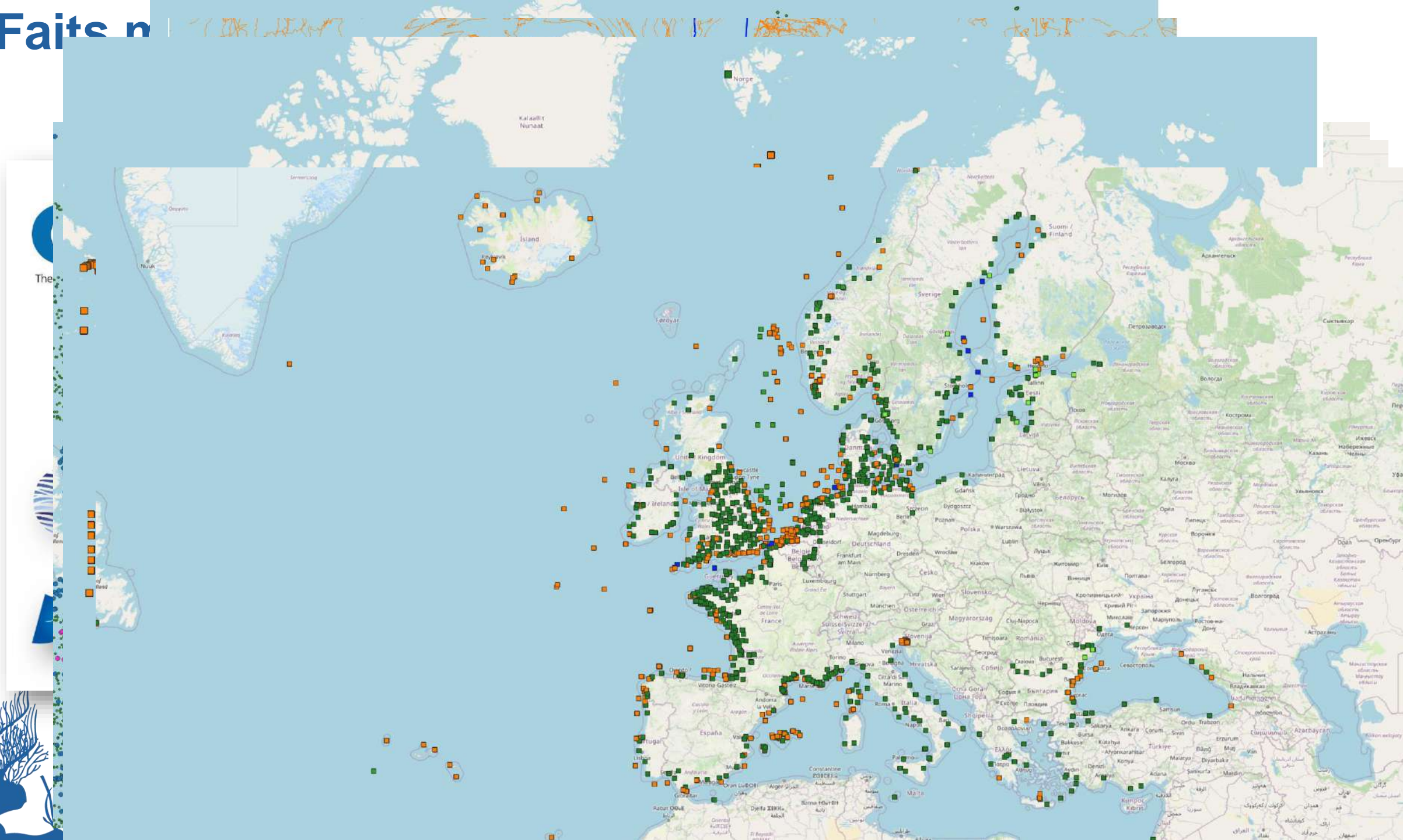
- EOSC-climat : UC Euro-Argo et ICOS ERICs flux de pCO₂ pour WMO GGGW
- Euro-Argo-One : vers une infra DAC-GDAC cloud

Echecs

- BG-Ocean échec de la proposition, sera resoumise
 - indicateurs sur les EOv et phénomènes Acidification, Deoxygenation, Carbone, Flux GES
 - initier un GDAC Oxygen (lead université de Liège)

EOSC-Ocre -> diffusion données Argo sur cloud Google

Faith n



The

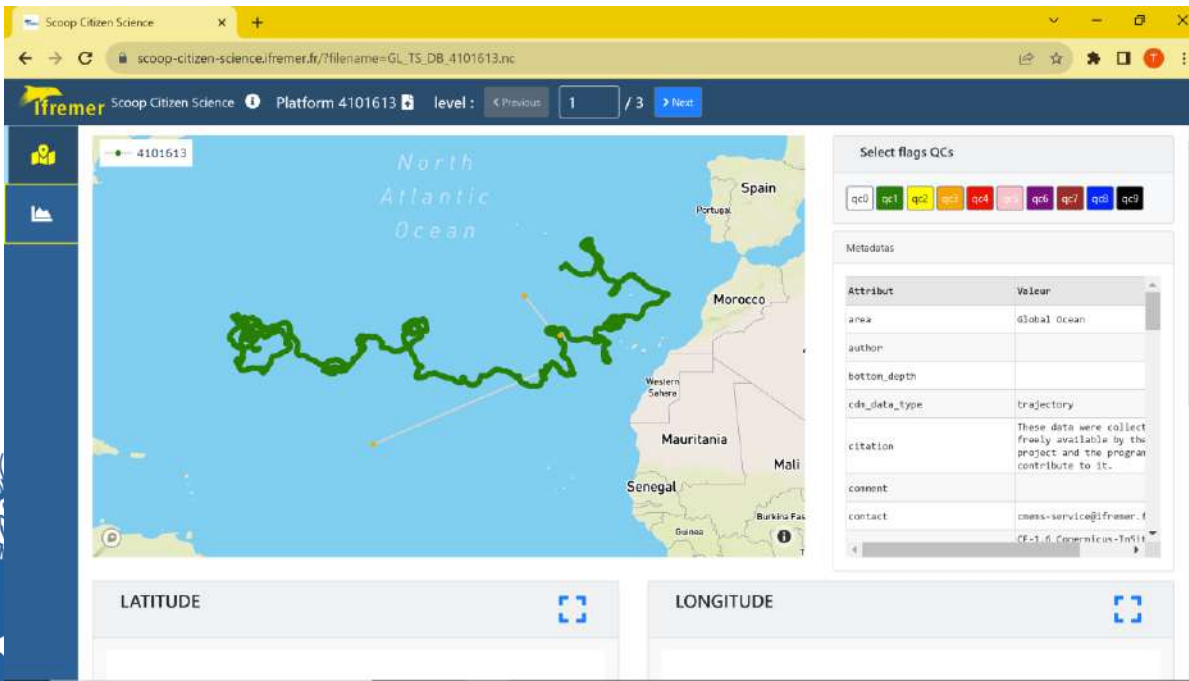
of

of



Perspectives

- Nouvelle version web CDS-Coriolis : éditorial + tableau de bord FR-OOS
- Flux BGC Cyber-Lefe vers Coriolis – Copernicus - Argo
Noyau de la base de référence pour BGC-Argo ? (données premium)
- Version Arco (zarr, parquet) du jeux bouées dérivantes
Bigdata bouées dérivantes, application à Scoop citizen science et IA

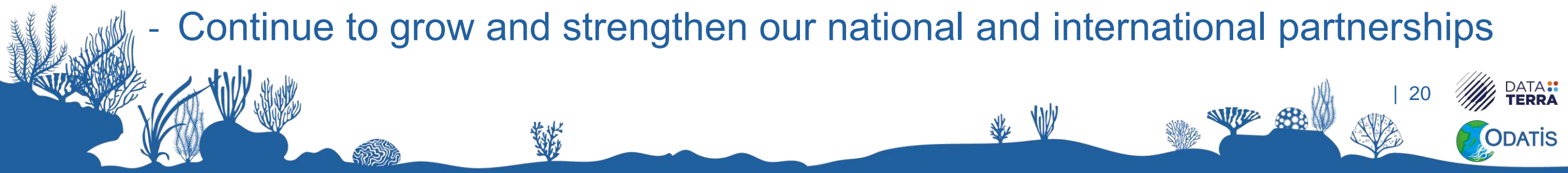


The screenshot displays the Scoop Citizen Science web interface. The main map shows the North Atlantic Ocean with a green trajectory line. The interface includes a sidebar with navigation icons, a top navigation bar with the ifremer logo and platform information, and a right-hand panel with a 'Select flags QCs' section and a 'Metadatas' table.

Attribut	Valeur
area	Global Ocean
author	
bottom_depth	
cds_data_type	trajectory
citation	These data were collect freely available by the project and the program contribute to it.
comment	
contact	ocean-service@ifremer.fr CE-1.6.Copernicus-Trajectory

Perspectives, OneArgo

- OneArgo : a new global design is under development
 - towards spatial completeness to include Polar sea-ice zones and marginal seas
 - increasing regional resolution in key areas like the Western Boundary currents and equatorial regions
 - major new missions: BioGeoChemical BGC-Argo and Deep-Argo missions
- Challenges
 - Significant new resources are needed – roughly 4-5 times the current investment
 - Requires strong community support and advocacy
 - Ensure our **data management teams are ready and able to handle the new data streams**
 - Continue to grow and strengthen our national and international partnerships



Perspectives, OneArgo

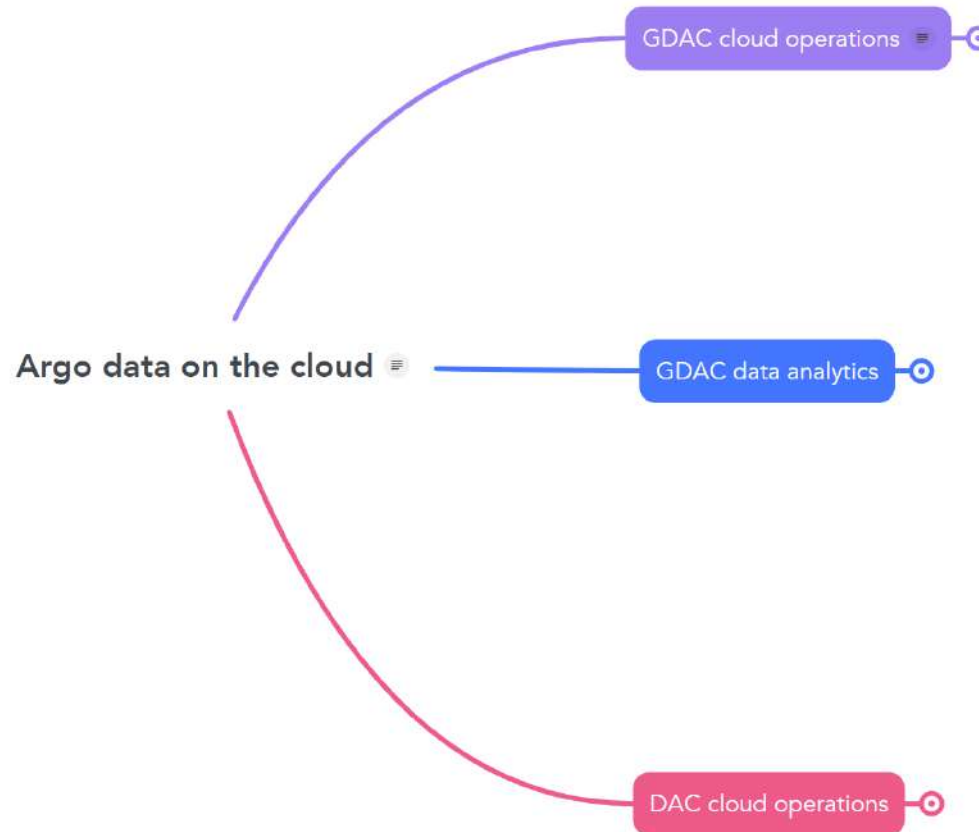
The Argo DAC-GDAC infrastructure, based on FTP transfers, was conceived a quarter of a century ago. A significant transition towards a cloud-based infrastructure is needed, driven by the following key objectives:

- **Scalability:** to accommodate the evolving demands of Argo data management and ensure seamless growth.
- **Reliability:** enhancing the trustworthiness and consistency of data operations, minimizing downtime.
- **Robustness:** strengthening the infrastructure's resilience and capacity to withstand challenges and disruptions.

This shift to a cloud-based approach will improve the efficiency of Argo data management and align it with contemporary technological standards.

Perspectives, OneArgo

- To have a clear idea of what is needed and where we go, here is a mind map listing the functionalities of **Argo cloud services**.



Besoins budgétaire et en personnel

- Budget

- Conforme à la convention inter organisme
- Partie Ifremer financée par projets Copernicus, ENVRI-hub-next, Blue-Cloud, EOSC-Future, EOSC-Fair-Ease, EMODnet-physics, Argo TGIR, Argo CPER ObsOcean

- Personnel (nouveautés)

- Ifremer

- Léo Le Lonquer en septembre, ingénieur cloud et bigdata
- Vanessa Tosello en novembre, ingénieur IA
- Ingénieur IA en cours de recrutements (206 candidats)
- Départ de Grégoire Mureau (CDD Copernicus Marine)



Attentes CDS-Coriolis vis-à-vis d'Odatis

- Promouvoir l' « EuroGOOS data policy »
 - Ethically open data: “as open as possible, as closed as necessary”
 - Clear, unambiguous data licenses such as CC-BY
 - Long term archives in certified repositories
 - FAIR principles
- Développer, enrichir les jeux de données de référence in situ
Crucial pour OneArgo (entre autres)