



DATA
TERRA



ODATIS

Atelier Technique « Accès aux données » Restitution du questionnaire

Caroline Mercier, Jean-François Piollé

10 octobre 2024 - Plouzané



Questionnaire

- Un questionnaire a été proposé aux CDS entre le 10/09 et 05/10 sur une page cachée du site ODATIS
- Réalisé avec Framiforms
- 53 questions ouvertes ou fermées, non obligatoires
- Modifications apportées aux questions pendant la phase d'ouverture du questionnaire pour mieux représenter les différentes filières d'un même CDS
- 7 participants

1. Stockage

1.1 Localisation des données de votre CDS

1.2 Accès externe

1.3 Accès interopérables - machine to machine

1.3.1 Services Opendap

1.3.2 Service S3

1.3.3 Services OGC: WCS, WMS

2. Catalogue et sélection de données

2.1 Catalogue ODATIS

2.2 Catalogue / Inventaire STAC

0. Préambule

Filière(s) de données concernée(s) par la réponse

Un même CDS peut être amené à soumettre plusieurs réponses. Il est par exemple recommandé de faire une réponse par **grande filière de données** si les réponses correspondant aux questions sont très dépendantes de celles-ci

3 réponses

- Données des SNO PHYTOBS et BenthOBS
- Biogéochimie des campagnes hauturières (LEFE-CYBER) + Boussole
- Données du CDS CERSAT dans leur ensemble

1. Stockage

1.1 Localisation des données de votre CDS

	L'ensemble de vos données référencées et décrites dans ODATIS est-il stocké dans votre CDS ?	Précisez notamment si ces données sont archivées par d'autres partenaires ODATIS, dans Seanoe, etc...
CDS-AVISO	Oui	
CDS-CERSAT	Oui	Non
CDS-OASU	Non	
CDS-Coriolis	Oui	
CDS-Shom	Oui	
CDS-STAMAR Roscoff	Oui	Des jeux de données annuels consolidés sont déposés sur SeaNOE
CDS-STAMAR Villefranche	Non	Le projet est qu'une fois que les données sont validées et ouvertes, elles sont envoyées sur SEANOE pour augmenter la facilité de citation

Quel mode de stockage est utilisé ?

	fichiers accessibles sur le réseau	fichiers sur disques amovibles	base de données	autre; précisez ci-dessous	Merci de préciser si besoin tout détail relatif au mode de stockage :
CDS-AVISO	X				
CDS-CERSAT	X				Tout est sur les disques de Datarmor, back-up des données critiques sur tape library
CDS-OASU	X		X		
CDS-Coriolis	X		X		
CDS-Shom			X		L'ensemble des données est disponible sur data.shom.fr
CDS-STAMAR Roscoff	X		X	X	Fichiers stockés dans l'infrastructure de stockage de la plateforme ABiMS (stockage mutualisé, sécurisé, sauvegardé).
CDS-STAMAR Villefranche	X				

Une partie de ces données est-elle recopiée et disponible sur une autre plateforme ou service ?

	Ifremer / Datarmor	CNES / TREX ou HAL	Seanoë	autre	Merci de préciser si besoin tout détail relatif sur ces plateformes ou services:
CDS-AVISO		X		X	Les données du CDS AVISO sont archivées sur le HPC du CNES (cluster TREX, copie primaire) et sur une plateforme à CLS pour les produits de niveaux 3 et 4.
CDS-CERSAT	X				
CDS-OASU			X	X	seadatanet
CDS-Coriolis	X		X		
CDS-Shom	X				Ifremer – Coriolis : collecte horaire des données temps réel de marégraphe data.shom et une collecte hebdomadaire des données temps différé de marégraphe data.shom. Données visibles sous dans data.selection (https://dataselection.coriolis.eu.org/) Les données d'observations du niveau de la mer sont également disponible sur divers portails européens ou internationaux (CMEMS, Emodnet Physics, SONEL, GLOSS,...)
CDS-STAMAR Roscoff			X		Une partie des données primaires des SNO BenthOBS & PHYTOBS est stockée dans Quadrigé, dont des extractions sont agrégées avec les données stockées en local pour construire les jeux de données annuels.
CDS-STAMAR Villefranche	X		X		Il est également prévu de servir une partie de ces données sur ERDDAP (ODATIS/ifremer)

1.2 Accès externe

Vos données ODATIS sont-elles accessibles aux utilisateurs ODATIS (autres que ceux de votre CDS ou laboratoire)?

Si Oui à la question précédente, via quel protocole d'accès ?

		Accès (type posix) local au sein du CDS, avec possibilité de SSH	Accès (type posix) local au sein d'une plateforme GAIA DATA*	FTP	HTTPS	Service web	Autre	Si vous avez répondu "Service web" ou "Autre" à la question sur les protocoles d'accès ci-dessus, merci de préciser
CDS-AVISO	Oui	X	X	X	X	X		Accessible via un serveur THREDDS et les services/protocoles associés (OPeNDAP, NetcdfSubset, HTTPServer,...)
CDS-CERSAT	Oui		X	X	X			
CDS-OASU	Oui			X	X			
CDS-Coriolis	Oui		compatible?	X	X	X		
CDS-Shom	Oui				X	X		API data.shom.fr : https://services.data.shom.fr/support/fr/services/refmar
CDS-STAMAR Roscoff	Oui				X		X	Via des applications de bancarisation dédiées et disponibles sous forme d'applications Web, à la communauté de producteurs CNRS/Universités.
CDS-STAMAR Villefranche	Oui				X			

(ex: Ifremer/Datarmor ou CNES/TREX pour ODATIS)



1. Stockage

1.2 Accès externe

Si vous avez répondu "Non", ou seulement en partie, à la question sur les protocoles d'accès; merci de préciser pourquoi ?				
ces données n'ont pas vocation à être distribuées	je n'ai pas eu le temps ou de moyens pour le faire	Autre, merci de préciser ci-dessous	Si vous avez répondu "Autre" à la question précédente, merci de préciser :	Commentaires ou questions libres concernant cette partie "Accès externe"

1. Stockage

1.3 Accès interopérable – machine to machine

1.3.1 Service Opendap

Vos données ou une partie de vos données sont-elles servies par OpenDAP?

	Servies par Opendap?	Si Oui à la question précédente, via quelle implémentation d'OpenDAP?					Si Non à la question "OpenDAP", pour quelle(s) raison(s) ?				Si vous avez répondu "Autre" à la question précédente, merci de préciser :
		ERDDAP	THREDDS	HYRAX	PyDAP	Autre	*	Mes données ne s'y prêtent pas	pas eu le temps de le faire	Autre	
CDS-AVISO	Oui		X								
CDS-CERSAT	Oui		X								
CDS-OASU	Non								X	X	Il n'a pas un besoin fort en interne de la part des chercheurs. Peut-être que « Build it, they will come » sera vrai.
CDS-Coriolis	Oui	X	X								
CDS-Shom	Non									X	Les APIs actuelles semblent satisfaire les besoins des utilisateurs.
CDS-STAMAR Roscoff	Non									X	À première vue, mes données ne s'y prêtent pas suffisamment pour proposer une interface de consultation/extraction basée sur OpenDAP.
CDS-STAMAR Villefranche	Oui	X					↑				

*Je dois déployer un serveur OpenDAP ou héberger mes données sur un site fournissant ce service.

1.3.1 Service Opendap

Si une partie seulement de vos données sont servies par OpenDAP, pour quelle(s) raison(s) ?

	Si une partie seulement de vos données sont servies par OpenDAP, pour quelle(s) raison(s) ?		
	Certaines ne s'y prêtent pas	Je n'ai pas eu le temps de le faire	Autre; merci de préciser ci-dessous
CDS-AVISO			
CDS-CERSAT		X	X
CDS-OASU			
CDS-Coriolis		X	
CDS-Shom			
CDS-STAMAR Roscoff			
CDS-STAMAR Villefranche		X	

1.3.1 Service Opendap

Difficultés, performances, intérêt à déployer ce service

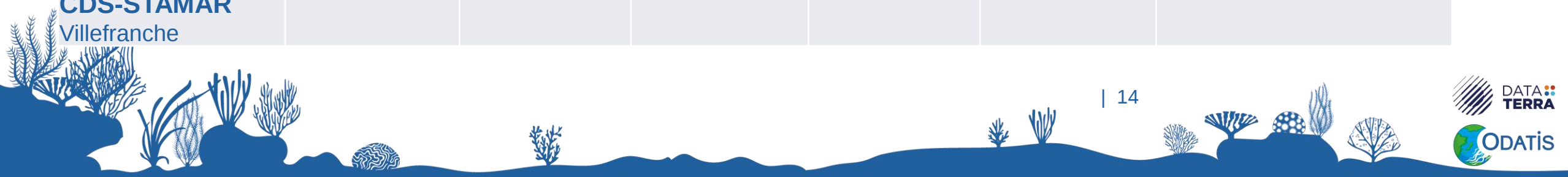
	difficultés particulières à déployer ou configurer un serveur OpenDAP?	problèmes de performance ou de maintenance avec votre serveur OpenDAP?	Seriez-vous intéressés pour déployer ce type de service ?	Commentaires ou questions libres concernant cette partie "OpenDAP"
CDS-AVISO	Difficulté de montée de version du TDS (v4.5 vers v5.5).			
CDS-CERSAT	manque de support pour configurer systématiquement les données sur un serveur OpenDAP. Problèmes de performance (cf-ci-dessous) n'encouragent pas à le faire.	solution THREDDS peu performante pour certains gros jeux de données (en particulier l'aggrégation). ERDDAP ne fonctionne pas sur nos données satellite. HYRAX devrait être évalué. Une solution stable et pérenne reste encore à définir.	Oui	
CDS-OASU				
CDS-Coriolis			Oui	
CDS-Shom			Je ne sais pas	
CDS-STAMAR Roscoff				
CDS-STAMAR Villefranche			Oui	pour l'instant les données sur ERDDAP ne sont disponibles qu'au sein du réseau de Villefranche, je manque de temps pour tout finaliser correctement

1.3.2 Service S3

		Si "Non" à la question précédente, ou seulement en partie, pour quelle(s) raison(s) ?						
	Vos données (...) sont-elles servies via un stockage objet et protocole de type S3 ?	Je n'ai pas prévu de le faire car cela n'était pas identifié comme un besoin	Je dois déployer un stockage objet dans mon infrastructure ou héberger mes données sur un site fournissant ce service.	Mes données ne s'y prêtent pas	pas eu le temps de le faire	Autre	Si vous avez répondu "Autre" à la question précédente, merci de préciser :	Commentaire
CDS-AVISO	Oui							
CDS-CERSAT	Non		X					
CDS-OASU						X	En cours d'évaluation. Dépend de plusieurs facteurs dont choix de l'hébergeur physique (datacenter)	
CDS-Coriolis	Oui							
CDS-Shom	Non					X	Protocole en cours de tests sur certains projets qui concernent d'autres types de données. Pas envisagé à ce stade pour CDS Shom.	
CDS-STAMAR Roscoff	Non	X						
CDS-STAMAR Villefranche	Non	X						

1.3.3 Services OGC: WCS, WMS

	Si vos données ou une partie de vos données sont servies par des protocoles OGC; pouvez-vous préciser lesquels ?					
	WCS	WFS	WMS	WTS	Autre	Si vous avez répondu "Autre" ci-dessus, merci de préciser:
CDS-AVISO			X			
CDS-CERSAT			X			
CDS-OASU						
CDS-Coriolis	X	X	X			
CDS-Shom					X	API inspirée du protocole SOS mais adaptée pour les besoins de la donnée et de son management.
CDS-STAMAR Roscoff			X			
CDS-STAMAR Villefranche						



1.3.3 Services OGC: WCS, WMS (suite)

	Si vos données ne sont pas servies par protocoles OGC ou seulement en partie, pour quelle(s) raison(s) ?						
	Je n'ai pas prévu de le faire car pas identifié comme un besoin	Je ne connais pas les solutions techniques permettant d'offrir ces services	Mes données ne s'y prêtent pas	pas eu le temps ou les moyens de le faire	Autre	Si vous avez répondu "Autre" à la question précédente	Commentaire
CDS-AVISO							
CDS-CERSAT		X				WMS repose principalement pour nous sur le serveur THREDDS dont les problèmes de performance et de configuration ont été mentionnés ci-dessus	
CDS-OASU					X	Certaines données le seront. Pour l'instant c'est en recette, donc pas officiel. Pour d'autres données, il n'y a pas de demande.	
CDS-Coriolis							
CDS-Shom							
CDS-STAMAR Roscoff							
CDS-STAMAR Villefranche				X			

2. Catalogue et sélection de données

2.1 Catalogue ODATIS

		Si "Non" à la question précédente, pourquoi ?			Pour les données disponibles dans le catalogue ODATIS, l'ensemble des moyens d'accès décrits ci-dessus (sections 1.2 et 1.3) y est-il documenté et à jour ? Si Non, pourquoi:			
	L'ensemble de vos données ODATIS sont-elles visibles dans le catalogue ODATIS (Sextant) ?	Reste encore des informations à saisir et pas eu le temps ou les moyens	Autre	Si vous avez répondu "Autre" à la question précédente, merci de préciser	Il me reste encore des informations à saisir et je n'ai pas eu le temps ou les moyens	Autre	Si vous avez répondu "Autre"	Commentaire
CDS-AVISO	Oui							
CDS-CERSAT	Oui				X			
CDS-OASU	Oui							
CDS-Coriolis	Oui							
CDS-Shom	Oui			Précision CDS-IS-Shom : Périmètre = données d'observations du niveau de la mer Données hydrologiques via le CDS-IS-CORIOLIS		X	Les fiches sont à mettre à jour	
CDS-STAMAR Roscoff	Oui				X			
CDS-STAMAR Villefranche	Non	X			X			

2. Catalogue et sélection de données

2.2 Catalogue / Inventaire STAC

	Quel est votre niveau de familiarité avec STAC ?					Si vous avez répondu "Non" à la question précédente; merci de préciser pourquoi ?				
	Je n'en ai jamais entendu parler	J'en ai vaguement entendu parler chez mon coiffeur	Je vois de quoi il s'agit	Je l'ai déjà utilisé	Avez-vous testé l'intégration de vos jeux de données sous un catalogue STAC ?	pas eu le temps ou les moyens	Je n'y comprends rien ou manque d'expertise	Mes données ne s'y prêtent pas	Autre	Si vous avez répondu une réponse ci-dessus demandant précision; merci de saisir ces détails
CDS-AVISO			X		Oui					
CDS-CERSAT				X	Oui					
CDS-OASU	X				Non				X	Jamais entendu parlé. À tester.
CDS-Coriolis			X		Oui					
CDS-Shom	X				Non				X	
CDS-STAMAR Roscoff		X			Non		X			
CDS-STAMAR Villefranche		X			Non	X				

2. Catalogue et sélection de données

2.2 Catalogue / Inventaire STAC (suite)

	Si "Oui", quels types de données avez-vous configuré ?	Si "Oui", quelles difficultés ou limitations avez-vous rencontrées ?	L'ensemble de vos données vous semble-il intégrable dans un catalogue STAC ?	Pouvez vous gérer l'ensemble de votre inventaire de données au sein d'un catalogue STAC statique ?	Si un catalogue STAC dynamique (API) est nécessaire (cas de collections de nombreux fichiers), sur quels types d'index s'appuie-t-il (base de données, Elasticsearch, ...) ?	Comment comptez-vous implémenter votre API STAC ?	Commentaire
CDS-AVISO	Fichiers SWOT de niveaux L1 et L2 (archivage principalement).		Oui	Non	Elasticsearch principalement.		
CDS-CERSAT	principalement des granules de type L2 pour lesquels la recherche spatio-temporelle est un service nécessaire.	principalement le manque de solution technique "clé en main". Des services spécifiques doivent être implémentés pour l'indexation spatio-temporelle de nos données d'une part (Elasticsearch) et y brancher une API STAC. Travail toujours en cours.	Oui	Non	Elasticsearch. Approche avec GeoParquet également en test.	Implémentation "maison" d'un serveur STAC branché sur Elasticsearch	
CDS-OASU							
CDS-Coriolis	On commence par le jeu de données Argo	C'est une activité très récente, encore en mode R&D	Oui	Je ne sais pas	Pas de choix pour le moment, probablement duckdb	Travail en cours	On espère aboutir sous un mois ou deux pour Argo
CDS-Shom			Je ne sais pas	Je ne sais pas			
CDS-STAMAR Roscoff			Je ne sais pas	Je ne sais pas			
CDS-STAMAR Villefranche			Je ne sais pas	Je ne sais pas			