



Plan de gestion des données du projet “RIOMAR”

Plan de gestion de données créé à l'aide de DMP OPIDoR, basé sur le modèle " Agence nationale de la recherche (ANR): ANR - Modèle de PGD (français)" fourni par l'ANR.

Version 1

Action : Programme Prioritaire de Recherche « Un océan de solutions »

Acronyme du Projet : RiOMar

Durée du Projet : 72 mois (du 01/10/2022 au 30/09/2028)

Dernière date de révision du PGD : 17/04/2023

Table des matières

1. Documents de référence.....	3
2. Introduction.....	4
3. Contexte.....	4
4. Objet et périmètre du plan de gestion des données.....	6
5. Description des données et collecte ou réutilisation de données existantes.....	9
a. Comment de nouvelles données seront-elles recueillies ou produites et/ou comment des données préexistantes seront-elles réutilisées ?.....	9
b. Quelles données (types, formats et volumes par ex.) seront collectées ou produites ?.....	10
6. Documentation et qualité des données.....	11
a. Quelles métadonnées et quelle documentation (par exemple méthodologie de collecte et mode d'organisation des données) accompagneront les données ?.....	11
b. Quelles mesures de contrôle de la qualité des données seront mises en œuvre ?.....	11
7. Stockage et sauvegarde pendant le processus de recherche.....	12
a. Comment les données et les métadonnées seront-elles stockées et sauvegardées tout au long du processus de recherche ?.....	12
b. Comment la sécurité des données et la protection des données sensibles seront-elles assurées tout au long du processus de recherche ?.....	12
8. Exigences légales et éthiques, codes de conduite.....	12
a. Si des données à caractère personnel sont traitées, comment le respect des dispositions de la législation sur les données à caractère personnel et sur la sécurité des données sera-t-il assuré ?....	12
b. Comment les autres questions juridiques, comme la titularité ou les droits de propriété intellectuelle sur les données, seront-elles abordées ? Quelle est la législation applicable en la matière ?.....	13
c. Comment les éventuelles questions éthiques seront-elles prises en compte, les codes déontologiques respectés ?.....	13
9. Partage des données et conservation à long terme.....	13
a. Comment et quand les données seront-elles partagées ? Y-a-t-il des restrictions au partage des données ou des raisons de définir un embargo ?.....	13
b. Comment les données à conserver seront-elles sélectionnées et où seront-elles préservées sur le long terme (par ex. un entrepôt de données ou une archive) ?.....	14
c. Quelles méthodes ou quels outils logiciels seront nécessaires pour accéder et utiliser les données ?.....	14
d. Comment l'attribution d'un identifiant unique et pérenne (comme le DOI) sera-t-elle assurée pour chaque jeu de données ?.....	14
10. Responsabilités et ressources en matière de gestion des données.....	15
a. Qui (par exemple rôle, position et institution de rattachement) sera responsable de la gestion des données (c'est-à-dire le gestionnaire des données) ?.....	15
b. Quelles seront les ressources (budget et temps alloués) dédiées à la gestion des données permettant de s'assurer que les données seront FAIR (Facile à trouver, Accessible, Interopérable, Réutilisable) ?.....	15

1. DOCUMENTS DE RÉFÉRENCE

[1] Convention attributive d'aide n°ANR-22-POCE-0006

[2] Agence nationale de la recherche (ANR) : ANR - Modèle de PGD (français) - Modèle créé avec *DMP OPIDoR*.

Glossaire

Acronyme	Nom complet
ANR	Agence Nationale de la Recherche
PGD	Plan de Gestion des Données
GT PGD	Groupe de Travail Plan de Gestion des Données
WP	Work Package
AI	Artificial Intelligence
SNO	Service National d'Observation
IR-ILICO	Infrastructure de Recherche Littorale et Côtière
JERICO S3-DS	Joint European Research Infrastructure of Coastal Observatories
EMODnet	European Marine Observation and Data network
ODATIS	pôle de données et de services pour l'océan
DATA TERRA	centre de référence thématiques système terre et environnement
SEANOE	sea scientific open data edition
CMEMS	service de surveillance de l'environnement marin de copernicus
CeDONA	Centre de Données pour l'Observation en Nouvelle-Aquitaine
OASU	Observatoire Aquitain des Sciences de l'Univers
SISMER	Systèmes d'Informations Scientifiques pour la Mer
DOI	Digital Object Identifier
ARCHIMER	Archive institutionnelle de l'Ifremer
HAL	archive ouverte

2. INTRODUCTION

Ce document constitue le plan de gestion des données (PGD) du projet RiOMar (01/10/2022 – 30/09/2028). Il s'appuie sur le modèle fourni par l'ANR [2]. Ce PGD fournit une description de tous les types de données qui seront collectées et utilisées et de leur utilisation tout au long du cycle de vie du projet.

Il doit être mis à jour tous les 2 ans à compter de la signature de la convention attributive d'aide. Une version révisée est donc prévue en juillet 2024, juillet 2026 et juillet 2028. Chaque révision du PGD aura un numéro de version inclus dans la section administrative ci-dessus.

Une version mise à jour à la date de fin de projet sera fournie, en septembre 2028.

Les dispositifs décrits dans ce document sont décidés et mis en application au niveau du comité scientifique exécutif du projet.

3. CONTEXTE

Le projet RiOMar a pour objectifs de :

- (i) Augmenter le nombre des paramètres observés, l'échelle spatiale et temporelle des données océanographiques et écosystémiques pour générer des observations augmentées ;
- (ii) Développer des modèles pour le présent et le futur afin de construire le jumeau numérique de l'océan côtier et caractériser l'exposome des écosystèmes dans le présent et le futur, en relation avec l'intelligence artificielle ;
- (iii) Co-construire avec les gestionnaires de l'environnement des solutions fondées sur la connaissance scientifique du fonctionnement des milieux.

en se concentrant sur 5 zones côtières sous l'influence des fleuves (RiOMar) de France métropolitaine couvrant la Manche, l'Atlantique et la Méditerranée.

Dans le cadre de ce projet seront réalisés :

- (i) des observations augmentées au sein des 5 RiOMar étudiées dans le Projet ;
- (ii) Des simulations numériques menées sur les trois côtes intégrant les contraintes climatiques et anthropiques sur trois périodes du 21^{ème} siècle : le passé récent, le milieu du 21^e siècle et la fin du 21^{ème} siècle ;
- (iii) Une démarche de co-construction avec les gestionnaires de l'environnement pour une évaluation adéquate de leurs écosystèmes régionaux, apportant ainsi des solutions scientifiquement informées pour faire face à la crise environnementale côtière.

Ci-dessous, une vue schématique de l'organisation du projet RIOMAR et des différents groupes de travail (WPs) ainsi que des partenaires du projet RiOMar.

Figure 1 : Diagramme organisationnel du projet RiOMar

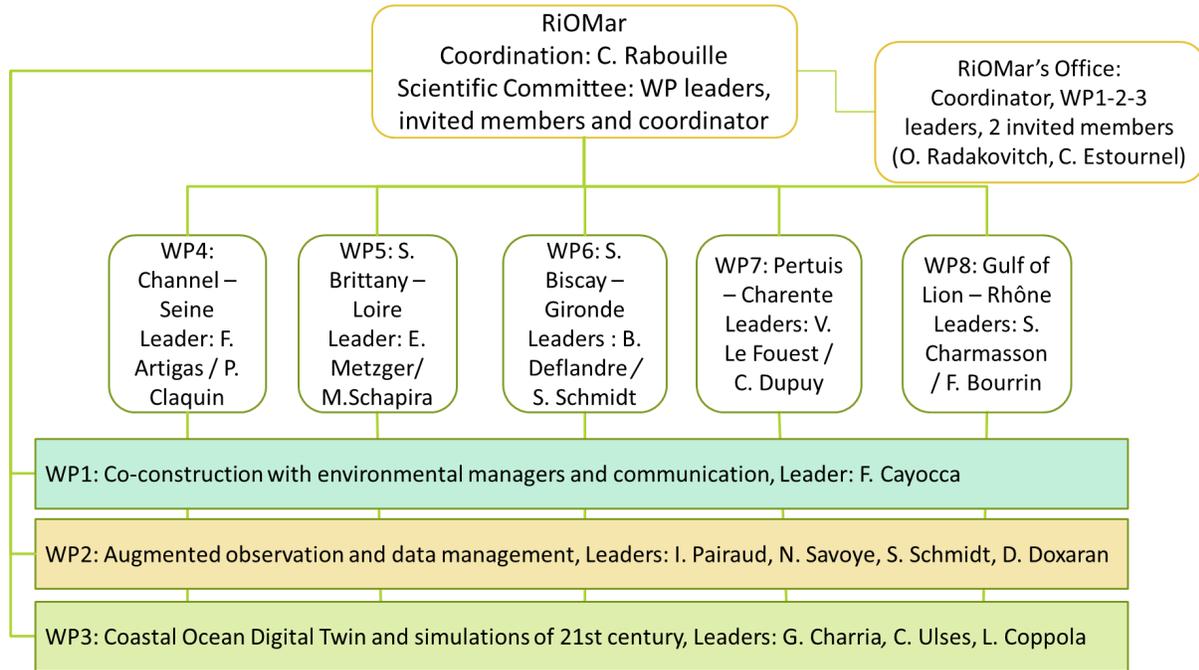
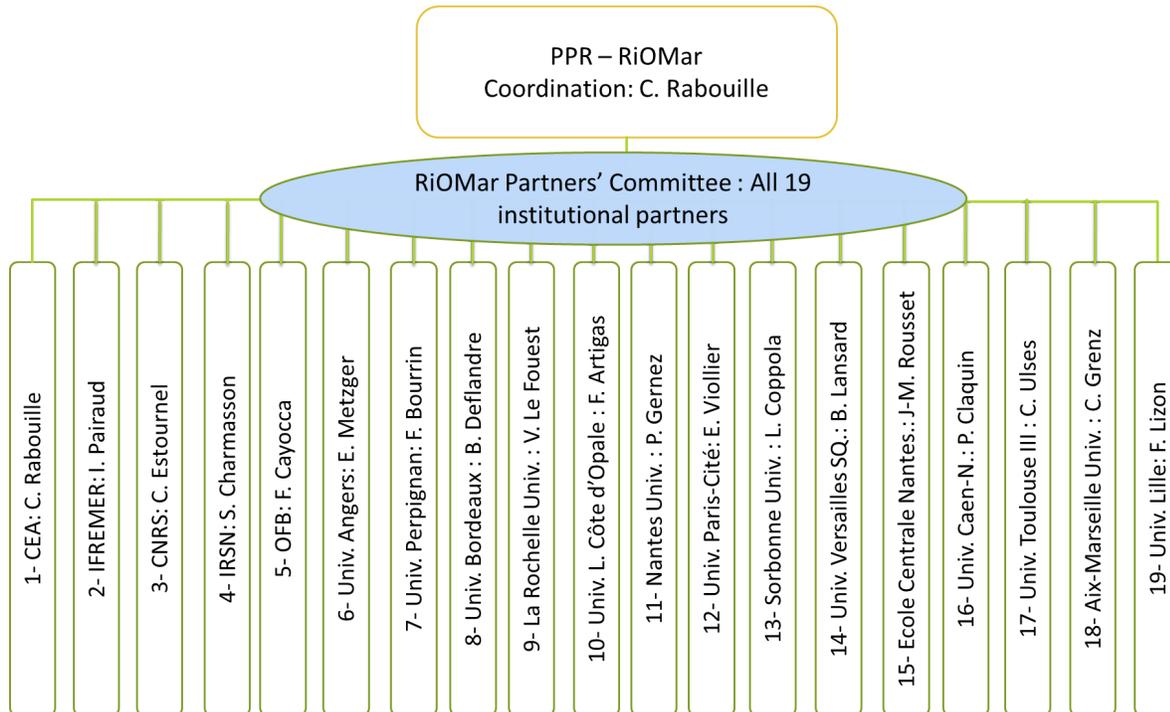


Figure 2 : Comité des partenaires de RiOMar



Au vu des différents objectifs et des réalisations prévues dans le cadre du projet, les données produites et réutilisées par les différents WPs et traitées dans ce document regroupent donc :

1. Les données numériques produites par les capteurs utilisés dans le cadre du projet, enrichies de leurs métadonnées (navigation, calibration, etc) ;
2. Les données numériques liées aux échantillons prélevés par les équipements utilisés dans le cadre du projet (par ex : données liées aux carottes sédimentaires, données de la colonne d'eau, etc) ;
3. La matière grise produite dans le cadre du projet : actes de colloques, rapports techniques à intérêt général, rapports de campagne, etc ;
4. Les publications à comité de lecture;
5. Les données numériques de modélisation et produites par l'intelligence artificielle et leur métadonnées (run de modèles, traitement ...) ;
6. Les données numériques satellites de couleur de l'océan.

4. OBJET ET PÉRIMÈTRE DU PLAN DE GESTION DES DONNÉES

Le présent plan de gestion de données, dans sa "version 1", ne constitue qu'une ébauche de l'inventaire exhaustif et détaillé des produits de recherche qui seront élaborés dans le cadre du projet RiOMar. Notre démarche pour accompagner la mise en place du PGD est la création d'un groupe de travail "donnée" (GT PGD) au sein de RiOMar qui regroupe un représentant par WP. Ce groupe se réunira autant de fois que nécessaire en vue de la révision du présent Plan de Gestion des Données.

Selon la diversité des données collectées et de l'étendue des chantiers géographiques, le nombre de représentant.e.s de chaque WP peut varier.

Contributeur / Contributeur	Affiliation	Rôles
Florence Cayocca	Office Français de la Biodiversité	Lien au WP1
Hélène Lecomte	Laboratoire des Sciences du Climat et de l'Environnement (CNRS, CEA, UVSQ)	Lien au WP1 et coordination
Christophe Rabouille	Laboratoire des Sciences du Climat et de l'Environnement (CNRS, CEA, UVSQ)	Lien aux WP1, WP8 et coordination
Sabine Schmidt	Laboratoire Environnements et Paléoenvironnements Océaniques et Continentaux (CNRS, UB)	Lien aux WP2 et WP6
David Doxaran	Laboratoire d'Océanographie de Villefranche (CNRS, SU)	Lien au WP2 (images satellites)
Joël Sudre	Infrastructure de recherche data terra	Représentant pour Data Terra

Caroline Ulses	Laboratoire d'Etudes en Géophysique et Océanographie Spatiales (UT, CNRS)	Lien au WP3 (modèle)
Laurent Coppola	Laboratoire d'Océanographie de Villefranche (CNRS, SU)	Lien au WP3 (IA)
Jean-François Leroux	Laboratoire de physique océanique et de télédétection par satellite (Ifremer)	Lien au WP3
Felipe Artigas	Laboratoire d'Océanologie et de Géosciences (CNRS, UL, ULCO)	Lien au WP4
Xavier Mériaux	Laboratoire d'Océanologie et de Géosciences (CNRS, UL, ULCO)	Lien au WP4
Michael Retho	Unité de recherche LITTORAL (Ifremer)	Lien au WP5
Vincent Le Fouest	Littoral ENvironnement et Sociétés (CNRS, UR)	Lien au WP7
Thibault Coulombier	Littoral ENvironnement et Sociétés (CNRS, UR)	Lien au WP7
Théo Berguig	Littoral ENvironnement et Sociétés (CNRS, UR)	Lien au WP7
François Bourrin	Centre de Formation et de Recherche sur les environnements méditerranéens (CNRS, UPVD)	Lien au WP8

Comme précisé en [3], ce PGD constitue un document de référence pour la centralisation et le stockage des différents types de données précisés plus haut, ainsi que pour la mise en place des métadonnées associées.

Une nomenclature permettant le tri des produits de recherche sera mise en place par le GT PGD. Elle décrira notamment le type de produit : jeux de données, produit de modèle, document de synthèse, etc.

Ci-dessous, la liste de l'ensemble des produits de recherche attendus dans le cadre de RiOMar :

D1.1 Documents de recommandations co-rédigés par des chercheurs et des gestionnaires, adaptés aux collectivités territoriales françaises : DIRM, DREAL, Agences de l'eau. Résultats à partager avec le ministère chargé de la mise en œuvre des directives. (Y1-Y2)

D1.2 Recommandations et documents de synthèse co-rédigés par les chercheurs et les gestionnaires ; ces documents seront diffusés plus largement aux gestionnaires régionaux, aux administrations régionales et aux autorités nationales. (Y1-Y2)

D1.5 Conception et réalisation de sites web. Projets avec des classes. Produits pour les étudiants de l'enseignement supérieur, communication sur les réseaux sociaux. Rapports de stage (Y1-Y6)

D 2.1 : Développement de produits satellitaires adaptés aux zones RiOMar et interface (Y2-v1,Y5-v2).

D 2.2 : Un ensemble de transects de planeurs obtenus pendant les périodes d'observation spéciales dans la S. Biscaye et le Golfe du Lion (Y5).

D 2.3 : Rapport sur l'utilisation des systèmes d'observation à faible coût : forces et faiblesses (Y5).
D 2.4 : Rapport sur l'observation intelligente : premiers résultats et besoins futurs (Y5).
D 2.5 : Rapport sur l'identification des niveaux de contaminants métalliques dans les sédiments et les Kd des zones RiOMar (Y5).

D2.6a : Plan de gestion des données (Y1) ;
D 2.6b : Données FAIR disponibles sur le portail ODATIS (Y3-Y6).
D 2.7 : Rapport sur la conception d'un futur réseau d'observation intégré dédié au suivi des changements pluri-décennaux dans les zones RiOMar. (Y6)
D 2.8 : Un ensemble de données d'observation participative dans la zone de la Bretagne S. sur 2 ans (Y5).

D3.1 (Y1) : Rapport technique décrivant le protocole adopté.
D3.2a (Y2) : Simulations hydrodynamiques pour le climat actuel. **D3.2b** (Y2) : Simulations hydrodynamiques pour le climat futur. **D3.2c** (Y2) : Simulations couplées de forçage à l'échelle du bassin médian pour le climat actuel et futur.
D3.3 (Y5) : Simulations couplées pour le climat actuel et futur.
D3.4 (Y5-Y6) : Publication scientifique, rapport CMEMS des produits de l'IA et un ensemble de données prédites par les réseaux neuronaux stockées dans un dépôt de données (par exemple SEAONE).
D3.5 (Y3) : Publication scientifique ou rapport sur l'évaluation du modèle.
D3.6 (Y5,Y6) : Publications ou rapports scientifiques sur les évolutions simulées et les impacts des multi-stresseurs sur les organismes marins.
D3.7 (Y5) : Ensemble de données fusionnées basées sur des simulations et des observations (in situ et télédétections).
D3.8 (Y5) : Protocoles pour la génération d'indicateurs et indicateurs produits.

D4.1 (Y3) Analyse spatialisée du contexte hydrodynamique et de la connectivité de la Baie de Seine avec la Manche Est.
D4.2 (Y4) Une analyse temporelle et spatiale multi-échelle des principaux modèles d'abondance, de biomasse, de dynamique, de composition en termes de classe de taille et de groupes optiquement caractérisés du phytoplancton et du développement des efflorescences.
D4.3 (Y4) Analyse des caractéristiques photophysologiques du phytoplancton et estimation de la productivité primaire.
D4.4 (Y5) : Intégration de toutes les données in situ/in vivo avec les produits de télédétection afin de paramétrer les modèles hydrologiques et biogéochimiques.

D5.1 Un ensemble de données pour T, S, O2, fluorescence, pendant 3 ans avec un accès en temps réel via le serveur de données (Y3).
D5.2a validation des données satellitaires sur la période du projet, en utilisant les capteurs HF et l'échantillonnage de terrain (Y3).
D5.2b Manuscrit de thèse ou article (Y4).
D5.2c rapport/article sur la combinaison de la modélisation et de l'observation : analyse des risques des événements extrêmes du phytoplancton en utilisant les données de modélisation (WP3) sur l'eutrophisation et l'hypoxie dans les prochaines décennies (2050 - 2100) (Y5).
D5.3a rapport/article sur un modèle conceptuel des cycles des éléments dans le plateau continental armoricain (Y4).
D5.3b Manuscrit de thèse ou article (Y4).
D5.3c rapport/article sur un modèle couplé benthique/pélagique (Y6).

D6.1.1 : conception de l'amarrage instrumenté haute fréquence. Rapport technique (Y1).
D6.1.2 : rapport de campagne et métadonnées d'échantillonnage (WP2) [campagnes de maintenance tous les 3 mois (Y1-Y5)]
D6.1.3 : rapport de campagne et cartes 2D des descripteurs physiques et biogéochimiques à chaque opération et métadonnées d'échantillonnage (WP2).
D6.2.1 : rapport de campagne (Y1-Y4) et métadonnées d'échantillonnage (WP2) pour l'échantillonnage de descripteurs biologiques et biogéochimiques ciblés dans la colonne d'eau et les sédiments.
D6.2.2 : Ensemble de données, 1 rapport ou article (Y5) sur l'état écologique de l'habitat benthique et ses fonctions (ex. minéralisation, bioturbation).
D6.3 Un rapport de campagne (Y3) sur la cartographie de la surface/couverture de boue du WGMP et son évolution, des cartes et 1 rapport ou article (Y4).

D7.1 : Ensemble de données d'un système de surveillance in situ HF des intrusions de GEW dans le PS (Y4).

D7.2 : Une base de données de mesures multi-paramétriques HF hébergée par l'OASU CeDONA en relation avec DATATERRA/ODATIS (Y4).

D7.3 : Une carte des zones marines socio-économiques potentiellement à risque en réponse à l'exposition fluviale des facteurs de stress environnementaux (Y5).

D8.1 : Rapport ou publication sur l'étude à long terme de divers paramètres par Mastodon-2D dans le GOL (Y5).

D8.2 : Rapport ou publications sur l'impact des événements extrêmes sur les sédiments et la colonne d'eau à l'embouchure du Rhône (Y5).

D8.3 : Rapport de campagne (Y3), et 1 rapport ou publication sur les flux benthiques des sédiments et leur impact sur la qualité de la colonne d'eau. (Y4).

A ce stade du projet, les réponses ou développement à un grand nombre de questions ne sont pas encore envisageables. Des révisions de ce PGD sont donc prévues ainsi qu'un travail concerté avec le GT PGD afin de suivre l'évolution des produits de données, leur traitement et stockage.

Les différents paragraphes suivants sont donc abordés d'un point de vue général. Ils seront précisés et déclinés par la suite pour chaque WP.

5. DESCRIPTION DES DONNÉES ET COLLECTE OU RÉUTILISATION DE DONNÉES EXISTANTES

a. Comment de nouvelles données seront-elles recueillies ou produites et/ou comment des données préexistantes seront-elles réutilisées ?

Des données préexistantes seront utilisées dans le cadre du projet en respectant les règles de confidentialité/utilisation propres à chaque jeu de données.

Ces données regroupent :

1. Des données d'observation des Services Nationaux d'Observations labellisés (SNO): **COAST HF, SOMLIT, MOOSE, PHYTOBS**. Ces données sont liées à l'infrastructure de recherche IR-ILICO dont le PGD des différents SNO est accessible ici : <https://www.ir-ilico.fr/?PlanGestionDonnees>

- Des données provenant de programmes internationaux : **JERICO S3-DS** dont les données d'observation sont stockées sur le catalogue **EMODnet** (European Marine Observation and Data Network).

Les données produites dans le cadre du projet peuvent être classées selon leur type, comme détaillé en [3].

En fonction du type de données, les méthodologies ainsi que le processus de collecte et de mise à disposition pourront évoluer.

Les délais de transmission aux bases de données seront précisés lors des révisions à venir de ce PGD.

b. Quelles données (types, formats et volumes par ex.) seront collectées ou produites ?

Comme précisé plus haut, ce plan de gestion des données considère :

- Les données numériques produites par les capteurs utilisés dans le cadre du projet, enrichies de leurs métadonnées (navigation, calibration, etc) ;
- Les données numériques liées aux échantillons prélevés par les équipements utilisés dans le cadre du projet (par ex : données liées aux carottes sédimentaires, données de la colonne d'eau, etc) ;
- La matière grise produite dans le cadre du projet : actes de colloques, rapports techniques à intérêt général, rapports de campagne, etc ;
- Les publications à comité de lecture;
- Les données numériques de modélisation et produites par l'intelligence artificielle et leur métadonnées (run de modèles, traitement ...) ;
- Les données numériques satellites de couleur de l'océan.

Si l'on exclut la matière grise et les publications, le tableau modèle ci-dessous recense les données produites ou collectées par les différents équipements. Il sera complété au fil des campagnes océanographiques et des travaux en cours.

Ce tableau générique sera décliné pour chaque WP du projet.

	Type de données	Format / logiciel d'accès	Volume
Equipement	par exemple numérique (bases de données, tableurs), textuel (documents), image, audio, vidéo, et/ou médias composites.	la manière selon laquelle les données sont codées pour le stockage, généralement reflétée par l'extension du nom de fichier (par exemple pdf, xls, doc, txt, ou rdf).	qui peuvent être exprimés en espace de stockage requis (octets), et/ou en quantités d'objets, de fichiers, de lignes, et colonnes.

Concernant la matière grise, cette dernière est stockée sur un drive interne au projet (owncloud), dont la sauvegarde est assurée par le LSCE (Laboratoire des Sciences du Climat et de l'Environnement). Ces données sont toutes accessibles à des formats standards (docx, . xlsx et . pptx.). Ces dernières ne devraient pas dépasser les quelques Go en fin de projet.

Pour ce qui est des publications, elles seront accessibles, après publication, dans les journaux concernés et accessibles sur Hal via leur dernière version libre de droit. Leur volume sera de quelques Mo.

6. DOCUMENTATION ET QUALITÉ DES DONNÉES

a. Quelles métadonnées et quelle documentation (par exemple méthodologie de collecte et mode d'organisation des données) accompagneront les données ?

Une fiche de donnée et de leurs métadonnées sera associée à chaque type de données. Elle précisera également les conventions de nommage des différents fichiers afin de retrouver sans ambiguïté les métadonnées associées aux données.

Les métadonnées génériques suivront les standards utilisés dans la communauté lorsqu'ils sont déjà définis (datation, navigation, paramètres d'environnement, ...) et on suivra les recommandations du pôle Odatis pour les conventions, formats et attributs.

Dans le cas de nouveaux types de données ou métadonnées vouées à servir la communauté, le pôle Odatis pourra être consulté afin de proposer des standards via des groupes d'experts constitués pour une thématique donnée et ingérer dans les serveurs de vocabulaire nationaux et internationaux les nouveaux types de données, avec leurs définitions, leur protocole d'obtention et la méthodologie de traitement de la nouvelle données associés.

	Métadonnées	Format
type de données	Indiquer quelles métadonnées seront fournies pour aider à la recherche et à l'identification des données. Indiquer quels standards de métadonnées seront utilisés (par exemple DDI, TEI, EML, MARC, CMDI).	format des métadonnées

b. Quelles mesures de contrôle de la qualité des données seront mises en œuvre ?

Les procédures de contrôle de qualité et d'attribution d'un code qualité aux données seront discutées par le groupe de travail constitué autour du plan de gestion des données et seront alignées aux standards d'ILICO.

7. STOCKAGE ET SAUVEGARDE PENDANT LE PROCESSUS DE RECHERCHE

a. Comment les données et les métadonnées seront-elles stockées et sauvegardées tout au long du processus de recherche ?

En fonction du type de données, le processus de stockage et de sauvegarde sera adapté.

1. Les données à intérêt pour la communauté (La matière grise (actes de colloques, rapports techniques à intérêt général, ...), les publications) sont stockées et sauvegardées sur le cloud du projet (cloud du LSCE).

Le choix du cloud du LSCE permet de s'appuyer sur les outils de l'institut et de s'assurer des bonnes pratiques en termes de stockage et de sauvegarde.

2. Les données collectées pendant les missions en mer sont transmises aux centres de données et de service qui auront été identifiés (par exemple, le SISMER, SEANOE, ...). La description de la procédure de stockage et sauvegarde des données pendant la phase d'exploitation sera détaillée au fur et à mesure de l'avancement du projet. La procédure d'intégration des données collectées sera mise en place en concertation avec le pôle Odatis et ses centres de données et de services.
3. Le mode de stockage, de transmission et de sauvegarde des données de simulation et produites par l'intelligence artificielle sera actualisé dans les prochaines versions du PGD.

b. Comment la sécurité des données et la protection des données sensibles seront-elles assurées tout au long du processus de recherche ?

Les données générées au sein de RiOMar seront stockées sous embargo, lesquels seront discutés avec le groupe de travail constitué et le Comité Scientifique Exécutif. Dans un souci de science ouverte, la mise sous embargo des données sera limitée au temps nécessaire pour les partenaires du projet d'en tirer le bénéfice [voir 9.a].

La gestion des données s'appuiera autant que faire se peut sur le pôle Odatis qui fédère au niveau national des activités de gestion de données et d'expertise scientifique en océanographie, ainsi que sur les centres de données et de services existants, ce qui permet de garantir les protections requises pour le stockage et la sauvegarde des données.

Les cas d'exception peuvent donner lieu à des principes d'accès avec des modalités de contrôle adaptées. Comme détaillé en [8.a].

8. EXIGENCES LÉGALES ET ÉTHIQUES, CODES DE CONDUITE

a. Si des données à caractère personnel sont traitées, comment le respect des dispositions de la législation sur les données à caractère personnel et sur la sécurité des données sera-t-il assuré ?

Des données à caractères personnels seront traitées via l'intervention d'un sous-traitant au projet : l'association Astrolabe Expédition.

Dans le cadre de leur participation, des données (Nom, prénom) de plaisanciers seront conservées de manière interne à l'association. Ces données seront associées à un fichier de date / localisation et de relevés liés aux capteurs utilisés.

Les données nom / prénom ne seront pas accessibles, ni aux partenaires du projet, ni à l'établissement coordinateur. Ces dernières sont stockées sur un drive interne à l'association et accessibles uniquement à son président et la salariée en charge du projet RiOMar. Elles n'ont pas vocation à être diffusées.

Afin de garantir la protection des données par l'association Astrolabe expédition, il sera demandé aux plaisanciers participants aux projets de donner leur consentement éclairé concernant la préservation de leurs données personnelles.

La question de la diffusion du matériel photographique et vidéo produit dans le cadre du projet RiOMar s'alignera sur la réglementation de la RGPD.

b. Comment les autres questions juridiques, comme la titularité ou les droits de propriété intellectuelle sur les données, seront-elles abordées ? Quelle est la législation applicable en la matière ?

Concernant la question des droits liés à la publication des données. Cette question s'oriente entre la publication sous Licence Creative Commons, le dépôt sur Hal ou Archimer des dernières versions des articles accessibles et ce dans le souci de l'accessibilité des données et des standards de science ouverte.

Les aspects de propriété intellectuelle sont traités dans l'accord de consortium, dont la négociation est en cours.

Les cas particuliers seront traités à part.

Concernant les données provenant des SNO, la question de la réutilisation de ces données sera traité essentiellement avec l'IR ILICO.

L'utilisation de matériel photographique et vidéo produit au sein de RiOMar sera encadrée par l'utilisation d'un copyright.

c. Comment les éventuelles questions éthiques seront-elles prises en compte, les codes déontologiques respectés ?

Chacun des participants au projet est tenu de respecter les chartes d'éthique et de déontologie de son organisme d'affiliation dont les principes sont définis par la Charte nationale de déontologie des métiers de la recherche.

9. PARTAGE DES DONNÉES ET CONSERVATION À LONG TERME

a. Comment et quand les données seront-elles partagées ? Y-a-t-il des restrictions au partage des données ou des raisons de définir un embargo ?

Les publications seront déposées dans des archives ouvertes libres d'accès, HAL ou Archimer en fonction de l'institut de rattachement de l'auteur principal.

Les données issues des capteurs seront stockées et sauvegardées. Le partage à destination d'une communauté scientifique plus large se fera dans le cadre des bonnes pratiques de la science ouverte. un

embargo, dont la durée sera fixée par le Comité Scientifique exécutif et le groupe de travail PGD sera mis en place pour préserver l'accès préférentiel aux données par les équipes participantes au projet.

Les données issues des équipements utilisés dans le cadre du projet seront déposées dans des centres de données et de services du pôle Odatis, afin de permettre un partage pérenne.

La distribution se fera via le catalogue Odatis et le SISMER.

Les cas particuliers seront traités à part.

b. Comment les données à conserver seront-elles sélectionnées et où seront-elles préservées sur le long terme (par ex. un entrepôt de données ou une archive) ?

Ce point fera l'objet d'une révision lors des révisions de ce document, une fois que les premières données produites / collectées au sein du projet seront transmises aux bases de données adaptées.

c. Quelles méthodes ou quels outils logiciels seront nécessaires pour accéder et utiliser les données ?

Afin de mettre à disposition les données en suivant les bonnes pratiques de la science ouverte, il est prévu un partage des données, après embargo, sur des plateformes librement accessibles à la communauté scientifique.

Les données partagées seront accessibles sur internet au sein des différentes bases de données auxquelles il est fait référence dans ce document. Les formats des données seront précisés en 5.b.

d. Comment l'attribution d'un identifiant unique et pérenne (comme le DOI) sera-t-elle assurée pour chaque jeu de données ?

Un numéro d'identification unique fourni par les plateformes de sauvegarde, ou un DOI, sera attribué à tous les jeux de données numériques destinés à être sauvegardés et partagés.

Le tableau ci-dessous recense les différents types de données générées par le projet, le lieu de stockage visé et quel identifiant sera attribué. Pour le cas des campagnes en mer, ces dernières sont déjà répertoriées par le SISMER, avec un DOI.

Type de donnée	Entrepôt	Identifiant unique pérenne
Publications	HAL ou Archimer	DOI
Matière grise	HAL ou Archimer	DOI
Jeux de données	SISMER (ou autre Centre de Services) ou Pôle ODATIS DATATERRA	DOI

Données numériques liées aux échantillons	Pôle ODATIS DATATERRA	DOI
Données numériques des modèles et de l'intelligence artificielle	Pôle ODATIS DATATERRA et Copernicus/CMEMS	DOI
Données numériques satellite de couleurs de l'océan	Pôle ODATIS DATATERRA	DOI
Campagnes en mer	SISMER	DOI

Afin d'assurer la traçabilité du projet, il est envisagé d'attribuer un DOI au projet, via des fiches mises à disposition sur une archive ouverte. Les jeux de données produits, les échantillons, les publications et la matière grise y feront référence.

On s'attachera à automatiser dans la mesure du possible le lien entre les objets numériques. A l'heure actuelle, seul le lien entre les campagnes en mer et les publications est automatisé sur SEANOE.

10. RESPONSABILITÉS ET RESSOURCES EN MATIÈRE DE GESTION DES DONNÉES

a. Qui (par exemple rôle, position et institution de rattachement) sera responsable de la gestion des données (c'est-à-dire le gestionnaire des données) ?

Les animateurs des WPs ou une personne déléguée sont responsables de la gestion des données produites par leurs équipes respectives. Ils auront notamment pour mission de s'assurer que leurs équipes suivent des directives du PGD et de collecter les informations propres à la mise à jour de ce document. Un groupe de gestion des données composé de ces représentant.e.s a été constitué [voir 4.]. Leur rôle est d'assurer la révision de ce PGD ainsi que de renforcer la transmission des bonnes pratiques précisées au sein du PGD.

La mise en œuvre du PGD sera de la responsabilité du coordinateur du projet.

b. Quelles seront les ressources (budget et temps alloués) dédiées à la gestion des données permettant de s'assurer que les données seront FAIR (Facile à trouver, Accessible, Interopérable, Réutilisable) ?

Au sein du WP2 : observation augmentée et gestion des données, un CDD dédié à la gestion des données sera embauché (embauche prévue courant 2024-2025). Cette personne ressource, assistée par le GT PGD, permettra d'obtenir un appui technique pour la mise en place de données FAIR au sein du projet RiOMar ainsi que d'appuyer les équipes dans leur dépôt des données au sein d'archives.

La préparation des données en vue de leur partage sera assurée par les équipes en charge de la collecte de ces mêmes données avec l'aide éventuelle du CDD dédié aux données.

Concernant le stockage des données de simulation numérique, une budgétisation de ce stockage a été réalisée (estimation de 50 Tera Octets) par le CNRS à hauteur de 8000€.