

Compte-rendu de la réunion de lancement du CES Imagerie Optique Benthique du Pôle Océan ODATIS

13-14/02/2024, Ifremer, Plouzané/En ligne

1. Agenda

13 février - après-midi (14h - 17h)

- Présentation du Pôle ODATIS
- Présentation du CES et périmètre (donnée au sens large)
- Tour de table avec projets et activités en cours
- Présentation des résultats du sondage
- Présentation/définition des sujets pour les ateliers

14 février - matin (8h30 - 12h)

- 3 ateliers de 45 minutes par sujet :
 - Acquisition
 - Traitements
 - Organisation, gestion, archivage et dissémination des données
- « Vote » sur les activités listées dans les ateliers

14 février - après-midi (13h - 16h)

- Compte rendu des sous-groupes et échanges sur les actions prioritaires
- Pour la suite du CES : quelles activités avec qui et comment, pour quand ?



2. Participants

Ces deux journées ont rassemblé une trentaine de participants (31) dont 18 ont participé en présentiel et 13 en ligne. 10 organismes étaient représentés, répartis de la manière suivante :

Organisme	Nombre de participants	Type de participation
Ifremer (centres/stations : Bretagne, Méditerranée, La Réunion)	18	Présentiel/En ligne
Université de Bordeaux (UMR EPOC)	3	Présentiel/En ligne
MNHN (PNDB)	2	En ligne
Station Biologique de Roscoff	2	En ligne
CNRS (Data Terra)	1	Présentiel
Institut Méditerranéen d'Océanologie (MIO)	1	En ligne
ENIB (UMR Lab-STICC)	1	Présentiel
ISEN	1	Présentiel
LOCEAN	1	En ligne
MarBio Consulting	1	En ligne

3. Journée du 13 février

3.1. Présentation du Pôle ODATIS (Erwann Quimbert)

IR Data Terra = plusieurs pôles de service : Data océanique, terre solide, ... PNDB est le 5ème pôle

25 organismes partenaires et mise en place de dispositif global de données : partie ODATIS et Data Terra littoral et d'autres disciplines de la surface au fond

Enjeux : accès au data océan, fédérer la gestion de data et expertise scientifique

Pôle ODATIS : 8 organismes avec 9 centre de données et services (data in situ, et 2 satellites : AVISO (data altimétriques) géré par le CNES, CORIOLIS, Campagnes Océanographiques et produit de la base Q2, SBR (SNO, Phytobs et Benthobs)

ODATIS : 5 CES avec charte de consortium et formulaire de proposition

Offre de services ODATIS : accès calculateur Datarmor et Trex (CNES), Entrepôt de données, catalogage de data en s'appuyant sur Sextant, Visualisation interactive des données, EVR environnement virtuel de recherche avec Rstudio et Jupiter Hub, Ateliers techniques (11-13 MARS en Bioinformatique à Roscoff et atelier littoral)

Support : Projets en tant que porteur ou contributeurs (nationaux, internationaux), Lien avec les IR (Illico, Fr-OOS, Argo, Emzo), Activité éditoriale

3.2. Présentation du CES et périmètre (Catherine Borremans)

Historique de la création du CES : Expansion massive du volume de données de différents capteurs sur différents supports >> manque de standardisation des protocoles en plus de limitations de ressources humaines pour les exploiter >> besoin d'automatisation et de bases d'entraînements = verrous

Objectifs : communauté nationale utilisatrice de l'image pour étudier les écosystèmes benthiques >> Pôle ODATIS selon FAIR

Communauté : Recherche marine & computer Vision. Travail sur l'ensemble du Workflow de l'acquisition à la restitution et converger vers des formats, bonnes pratiques, archivage et accessibilité des données annotées pour les développements de l'IA

3.3. Tour de table

Florence Conquet (IFREMER/SISMER)

Service SISMER, objectifs du CES sont d'intérêt pour le SISMER plus spécifiquement pour l'organisation, archivage et dissémination des données car travaille sur l'outil EcoTaxa qui permet de traiter l'imagerie liée à l'étude du plancton dans le cadre du rapatriement de cette plateforme à IFREMER

C'est une application IA open source avec une partie offline développée par le LOV qui n'a plus les moyens d'assurer tous les aspects techniques et humains. C'est la plus grosse base de données mondiale et fonctionne avec du RF pour la classification

Claire Laguionie

Ecosystème benthique profond (taxinomiste) avec des images de ROV. Travaille pour le MNHN sur les holothuries. Catalogue d'espèces standardisé

Centre d'intérêt pour la standardisation (metadata, protocole et limite d'identification, échantillon ? méthode d'identification (ADN, spécimen, ...))

Matteo Contini (IFREMER/DOI)

Travaux de thèse qui intègrent beaucoup de data de 3 projets à la DOI (IOT, Seatizen et Plancha), avec des images drone

Protocole pour standardiser entre data brute et data traitées (IA et annotation)

Pas d'outils pour partager des données open source >> orienté vers Zenodo (permet d'automatiser l'upload des données (geoflow))

Besoin de voir si moyen de partager des data/informations autour des API et standardisation

Marc Hoebeke

Ingénieur à la Station Biologique de Roscoff

Travaux dans le cadre de la Plateforme ABIMS (stockage et BDD au service du SNO Illico, Phytobs et Benthobs, pas d'imagerie)

Data : Pagure, sédimentaire et EcoTaxa

Intérêt pour CES Imagerie car co-porteur de FUTUR-OBS, WP4 Data management

Besoin : Comment on peut fournir accès à ces data au format d'origine et en terme de metadata (protocole, acquisition et analyse, espèce et comptage d'espèce) ? et comment garder le lien entre data origine et produits ?

Séverine Martini (MIO)

Projet avec un robot BathyBot (EMSO) avec un observatoire câblé (depuis mai 2023) adossé à différents environnements. Différents instruments (2 caméras pour le suivi de la faune benthique et une ligne instrumentée (data en temps réel))

Types d'images : Récif artificiel 3D et suivi de sa colonisation et caméra à haute définition

Intérêt : Pas de schéma de traitement et besoin de rejoindre les travaux de IFREMER en sciences participatives et commencer à réfléchir aux annotations d'image pour la classification automatisée des réseaux de neurones des images de sciences participatives

Yvan Le Bras (MNHN)

Présentation du MNHN/PNDB

S'intègre dans un service national, impliqué dans Data Terra pour guider la communauté en termes de standards, le socio écosystème des data est très hétérogène

Propose des flux entre data, catalogues, entrepôts ...

Lien entre IR et Infrastructures internationales (GOOS) + occurrence d'espèces avec le travail de GBIF mondial mais aussi national

CES : Intérêt pour le volet partage, gestion et annotation d'images vidéo

Action nationale sur la gestion et l'annotation de média (EcoInfoFAIR 2024), plateforme pour mettre à disposition des data mais aussi des algorithmes, Plateforme Galaxy avec intégration de R Shiny, VIAME...

Melissa Hanafi-Portier

PostDoc (Taiwan) structure des assemblages de poisson (image ROV de sommet et pentes de sommets sous-marins)

Intérêt Image vidéo pour traitement d'image et faciliter le workflow entre logiciels (BIGLE, ArcGIS) et développement de standards pour identifier la faune

Besoin pour le substrat (peu de base de données) (Pourcentage de couverture ??, morphologie des éponges)

Centraliser les outils et améliorer les performances

Karine Olu (IFREMER/LEP)

Travail sur images de l'observatoire, annotation manuelle (chronophage avec BIIGLE)

Travail sur le substrat (important pour caractériser le benthos) >> hétérogène >> automatisation

Evolution méthode :

- Automatisation par le biais de l'IA pour la détection (collaboration avec ISEN)
- Standardisation des annotations au niveau national avec des limites d'identification avec quelle confiance

Lénaïck Menot/Marin Marcillat (IFREMER/LEP)

Data vidéo des engins sous-marins (ROV) 40 à 240 heures vidéo, camera fixe (100h/an)

Elles sont annotées avec BIIGLE (présente des typologies d'item faune (eg SMarTaR-ID), substrat (eg CATAMI), macro déchet) qu'on peut partager entre utilisateurs

MATISSE 3D en cours d'amélioration pour traiter de plus gros volumes de data, avec copies multiples (archivée SISMER, BIIGLE, ...) >> peut-on réduire ?

Utilisation de l'IA pour la reconnaissance de faune

Limitations : Nécessité de grosse puissance de calcul, pas d'archivage pérenne des annotations (pas accessible depuis BIIGLE) pas de lien avec le catalogue SISMER et également en termes de metadata

Discussion avec partenaires internationaux pour standardisation (eg IFDO), reprojection des annotations sur les modèle 3D pour plus de précision, annotation pour détection automatique de la faune

Alicia Romero (EPOC, Univ Bordeaux)

Travaille avec les écologistes benthiques à Bordeaux

Analyses diverses et objectifs différents :

- FUND (Filter feeders UNDER change) : comportement de nutrition des balanes,
- JERICO S3 : station de déploiement d'instrument avec une série d'images acquise plusieurs jours pour la bioturbation (SPI sédiment)
- FUTURE-OBS WP2 : développement pour analyse de séries temporelles avec suivi de traceurs et annotations sur image de Pagure de vidéo-transect (BIIGLE)

Antoine Lavrard et Théodore Boch (IFREMER/DYNECO)

En stage, estimation bioindicateurs biodiversité

Marjolaine Matabos (IFREMER/LEP)

Chercheur en écologie benthique

Intérêt pour la dynamique spatio-temporelle (7000 heures de vidéo en 10 ans)

Vu les évolutions des annotations (Espions des Grands Fonds) une solution pour traiter séries temporelles. La caméra bouge donc difficile d'annoter temporellement

Algorithmes de DL avec sous-traitance Capgemini pour pérenniser, reconstructions 3D

Intérêt pour les standards, y compris concernant l'acquisition (métadonnées), et une base de données d'annotation commune et de l'archivage ...

Clémence Cotten (IFREMER/SISMER)

SISMER (Pôle data biologiques de laboratoire)

Responsable de la FAIRisation des données (Pôle ODATIS)

Sollicitation pour module de gestion d'images dans LabCollector et Pôle de données campagne (base de données BIGOOD, intégration info sur échantillons à bord) et besoin de visualiser sous outils tels que Sextant

Besoins d'intégration data via photothèque, vidéothèque

Gwenaël Caer

CNRS, DataTerra (GaiaData) >> Faciliter l'accès au data mais aussi et VRE (Jupyterlab/hub, notebooks)

Travail sur cas d'usage du Pôle ODATIS >> traitement de données benthiques (comprendre le besoin pour implémenter dans une VRE)

Kamal Nasreddine (ENIB)

LABSTICC, IA pour la mer et l'environnement

Détection de poissons dans images vidéo sous-marine et automatisation de surveillance des comportements

Intérêt (projet BlueRevolution) pour la méiofaune >> Apprentissage profond pour des image 1D (ADN), 2D et 3D

Images synthétiques générées et fournies au système

Projets étudiants pour détecter des traits fonctionnels ou problèmes de sociétés (électronique, informatiques)

Olivier Fauvarque (IFREMER/RDT)

Compétence sur tous les capteurs photoniques (fluorimètre, interféromètre, imagerie microscopie en flux)

Projet KOSMOS (caméra rotative pour observer des habitats) besoin de revoir le workflow dans une perspective de sciences participatives pour l'identification des poissons et des habitats

Claire Rollet (IFREMER, Dinard)

Herbiers de zostères, benthos intertidal et petits fonds (Rebent) et cartographie au niveau plateau

Utilisateurs sur l'écran...

Besoins : gestion de gros volume de données, standardisation et metadata

Impliquée à l'échelle des façades de l'ortholittoral aux satellites mais besoin de comparer les supports

3.4. Présentation du questionnaire (Thibault Napoleon)

36 participants :

- Majoritairement ingénieurs et chercheurs
- Beaucoup d'instituts mais surreprésentation de l'IFREMER avec des domaines d'intervention différents (côtier, profond, halieutique, technologie et informatique)
- Plus de la moitié se pose la question d'acquisition des données
- Principalement photogrammétrie et vidéo mais aussi télédétection

Besoins :

- Traitement : cartographie d'habitats, analyse automatique, suivi temporel
- Types de données : données sous-marines et aériennes, satellite, metadata et standardisation des data
- Données spécifiques : habitats, évaluation de la diversité, écologie spatiale
- Autres : pas bien clair, besoin de réflexion pour les définir

Traitements : Matlab, ENVI, ImageJ, Pix4D, Matisse ...

Annotations : majoritairement oui mais 1/3 non. Outil qui sort du lot : BIIGLE

Analyse : R, QGIS, ...

Besoins : Voir slides

4. Journée du 14 février

4.1. Identification des besoins et actions prioritaires

BESOINS EN ACQUISITION

1. Catalogue de protocoles

Ex. banque de protocoles internationale : <https://www.protocols.io/>

https://repository.oceanbestpractices.org/bitstream/handle/11329/1782/SOP_iFDOCreation_v1.0.0.pdf?sequence=1&isAllowed=y

<https://zenodo.org/records/5683082>

2. Application des standards dès l'acquisition
 - Instrument en lien avec les datasets
 - Métadonnées (voir [iFDO](#))
 - Outils flexibles (ex. récupération automatique des metadata sur les vidéos)
 - Cahier des charges
3. Recommandations en termes d'acquisition d'images pour des débutants
4. Organiser un atelier sur les standards en lien avec les GT ODATIS « Instrumentations » et « Outils communs » (qui implique IR Flotte, NSE et SM)
5. Prise en compte de l'aspect participatif à l'acquisition (cf. collectifs existants et Data Terra), notamment en termes de recommandations

Protocoles et recommandations ne doivent pas forcément être standardisés au départ mais au moins centralisés et partagés.

BESOINS EN TRAITEMENTS

1. Prétraitements :
 - Comment prendre en compte l'IA : pas besoin de prétraiter, fournir les images brutes
 - Est-ce nécessaire et notamment dans une optique de partage des données ? >> cela peut dépendre de la méthode et algo pour extraire l'information thématique ; les images corrigées sont utiles pour l'annotation manuelle
2. Introduction du contexte dans l'algo d'apprentissage ?
3. Intégration des prétraitements en temps réel lors de l'acquisition (ex. turbidité) ? ex. caméras Forsea, Orphie, Paralenz (corrige complètement aspect bleu/vert) – peut être problématique pour l'IA
4. Catalogue de prétraitements en lien avec le GT ODATIS (création catalogue d'algorithmes)
5. Partage des codes sources identifiés par la communauté
 - A FAIRiser ? Gitlab ? [Software Heritage](#) ?
 - Besoin d'être accompagné sur le plan technique
 - Cadre d'initiatives existantes qui peuvent permettre de le faire (ex. EcoinfoFAIR)
6. Annotations :
 - Développement d'algos à intégrer dans les softs existants (ex. BIIGLE, CVAT)
 - Comment pérenniser ?
 - Intégrer dans des initiatives internationales (maintenir, mettre à jour librairies) ?
 - Quels outils ? Ecotaxa peut-il être une solution générique ?
 - Atomisation et généralisation à chaque étape du workflow
 - Science participative ? voir [biigleparty](#) ?

- Plugins à ajouter dans les solutions existantes (BIIGLE ?)
- Elaboration de bases d'annotations d'images partagées
- Exercice par « cas d'usage » >> cf. FATHOMNET ? (Nécessité de créer d'abord l'instance nationale avant de transférer vers l'international)
- S'appuyer sur de gros data centres connectés, cloud européen ?
- Qualifier la qualité des annotations (comment annoter correctement ?)
- Ne pas se limiter à 1 seul logiciel >> s'assurer de l'interopérabilité

BESOINS POUR L'ARCHIVAGE & LA DISSEMINATION

1. Utilisation des standards et normes (iFDO et correspondance avec OGC côté géospatial ou l'EML côté biodiversité / écologie ; CATAMI, deepseacatalogue ; vocabulaires BODC ?)
 - Flux de données (à préciser)
 - Volume de données
 - Ontologies pour décrire généalogie de la donnée
 - Outils
2. Quelle data archiver / bancariser
 - Filtre avant archivage ?
 - Annotations, modèles 3D, ... ?
3. Problème de duplication du stockage (datawork vs instance BIIGLE)
4. Quelle solution pour l'archivage des annotations ?
5. Faire l'exercice pour détailler le workflow (data et metadata) en lien avec le GT ODATIS-Flotte
 - Cf. projet CAMERA comme base de réflexion
 - Possibilité de « Use Case » dans le cadre de Gaia Data (Infrastructure nationale distribuée)
6. Dictionnaire terminologique pour l'ensemble du workflow depuis l'acquisition à la restitution
7. Organisation d'atelier pour décliner les étapes du workflow
 - A croiser avec les travaux des GT ODATIS
 - Lister inputs/outputs/metadata pour chaque étape
 - S'inspirer des travaux de GEOMAR
 - Intégrer la notion de précision/qualification des produits en sortie (ex. https://www.marine-imaging.com/fair/sops/SOP_image-curation/)

4.2. Vote sur les actions identifiées avec les participants

Numéro	Action CES ODATIS Imagerie Benthique	Vote
ACQUISITION		
1	Catalogue de protocoles	1
2	Définir les métadonnées d'acquisition (standards)	6
3	Recommandations pour les débutants	0
4	Gestion de l'acquisition participative	0
TRAITEMENTS		
5	Prise en compte des prétraitements ?	0
6	Catalogues de prétraitements et traitements en lien avec GT ODATIS	6
7	Partage de codes sources	0
8	Annotations : Outils et mise en œuvre pour la construction de bases d'images annotées partagées	12
9	Annotations : Choix des outils interopérables (incluant le participatif)	4
ARCHIVAGE & DISSIMINATION		
10	Organisation d'un atelier pour la définition du Data workflow de l'acquisition à la restitution	8
11	Elaboration d'un dictionnaire de terminologie	2
12	Solution pour l'archivage et l'accès aux données d'imageries (mosaïque, modèle 3D, annotations, ...)	6
13	Quels standards utiliser pour la gestion des données imagerie	9

En rouge : Actions qui ont été retenues majoritairement par les participants

4.3. Pour la suite du CES : quelles activités avec qui et comment, pour quand ?

- Organisation d'un atelier pour la définition du Data workflow de l'acquisition à la restitution
- Annotations : Outils et mise en œuvre pour la construction de bases d'images annotées partagées
- Quels standards utiliser pour la gestion des données imagerie