



**DATA
TERRA**



ODATIS

Atelier feuille de route du pôle ODATIS de l'IR Data Terra

Equipe de direction du pôle

Assemblée Générale ODATIS

12-13 décembre 2023

Observatoire Aquitain des Sciences de l'Univers, Pessac



15/12/2023

contact@odatis-ocean.fr | www.odatis-ocean.fr

Feuille de route

- Feuille de route intégrée (pôles, services transversaux) de l'IR Data Terra pour 2023
- Chaque pôle ou dispositif jusqu'au CDS dispose de feuilles de route
- L'IR Data Terra a sa feuille de route (FRN) mais elle reste une méta-FR
- Les activités de services proposées vont être de plus en plus intégrées et imbriquées (effet Gaia Data) et devoir disposer de modèles économiques

co-construire une feuille de route en 2024 avec IR Data Terra pour aval CODIR et mise en œuvre pour 2024/2025 - 2030

Programme d'actions 2020 – 2024 du pôle ODATIS

1. Proposer un ensemble de services techniques

- Elaborer un catalogue des données, des produits et des services
- Encourager l'utilisation des DOI
- Mettre à disposition des outils d'exploration (découverte via le catalogue), de visualisation et d'extraction
- Garantir la pérennité des jeux de données : s'adosser aux centres d'archivages nationaux (ex. cines), aux structures dédiées que sont les centres de données des Observatoires des Sciences de l'Univers (OSU) ou des organismes dotés de services spécialisés (CNES, Météo-France, Ifremer, Shom)
- Faciliter l'analyse et le traitement des données proposées en mettant en œuvre des plateformes informatiques virtualisées de traitement et d'analyse de données

Programme d'actions 2020 – 2024 du pôle ODATIS

1. Proposer un ensemble de services techniques
2. Renforcer l'interopérabilité des CDS
 - Interopérabilité sémantique : adopter ou développer des référentiels et vocabulaires communs (thésaurus, ontologies) permettant de décrire sans ambiguïté variables observées, unités, méthodes , etc. en s'appuyant sur les travaux des instances et programmes internationaux, européens ou nationaux
 - Interopérabilité technique : Des jeux de données harmonisés, tant sur les plans de la description, de la qualité, du format de diffusion, etc. en s'appuyant, quand ils existent, sur les normes préexistantes au sein des différentes communautés

Programme d'actions 2020 – 2024 du pôle ODATIS

1. Proposer un ensemble de services techniques
2. Renforcer l'interopérabilité des CDS
3. Débuter l'élaboration de produits à valeur ajoutée (CES)
 - Le pôle ODATIS recensera, dans un premier temps, les Produits générés par les Centres de Données et de Services (CDS). Ces Produits seront décrits et rendus accessibles par le Catalogue.

Dans un second temps, le pôle ODATIS incitera à la mise en place de Consortiums d'Expertises Scientifiques (CES) afin de définir et de prototyper de nouveaux Produits.

Programme d'actions 2020 – 2024 du pôle ODATIS

1. Proposer un ensemble de services techniques
2. Renforcer l'interopérabilité des CDS
3. Débuter l'élaboration de produits à valeur ajoutée (CES)
4. Assurer une communication efficace
 - Alimentation continue du portail web, newsletter, séminaires...

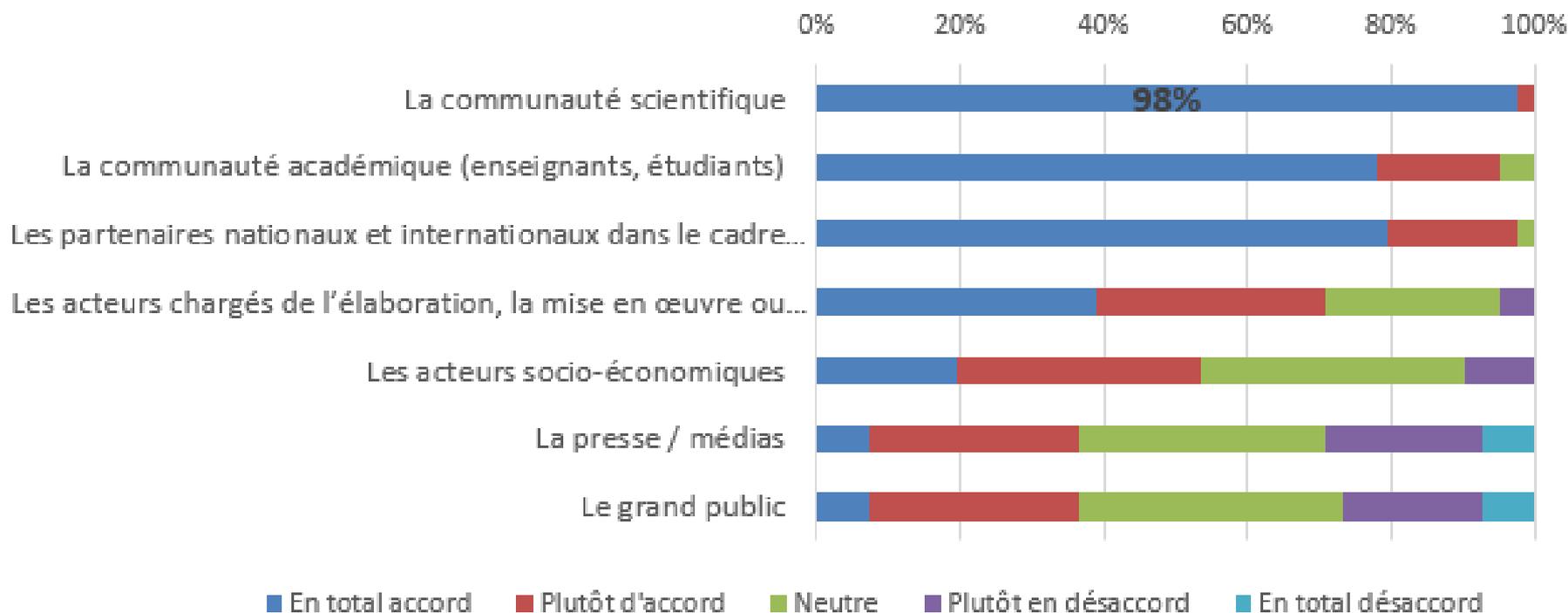
Programme d'actions 2020 – 2024 du pôle ODATIS

1. Proposer un ensemble de services techniques
2. Renforcer l'interopérabilité des CDS
3. Débuter l'élaboration de produits à valeur ajoutée (CES)
4. Assurer une communication efficace
5. Renforcer l'accompagnement utilisateurs
 - Dans un premier temps, l'effort d'accompagnement sera effectué au bénéfice des producteurs de données : Infrastructures de Recherche (par exemple, volets de l'IR Illico pour l'observation côtière), programmes et projets d'observation.

Questionnaire

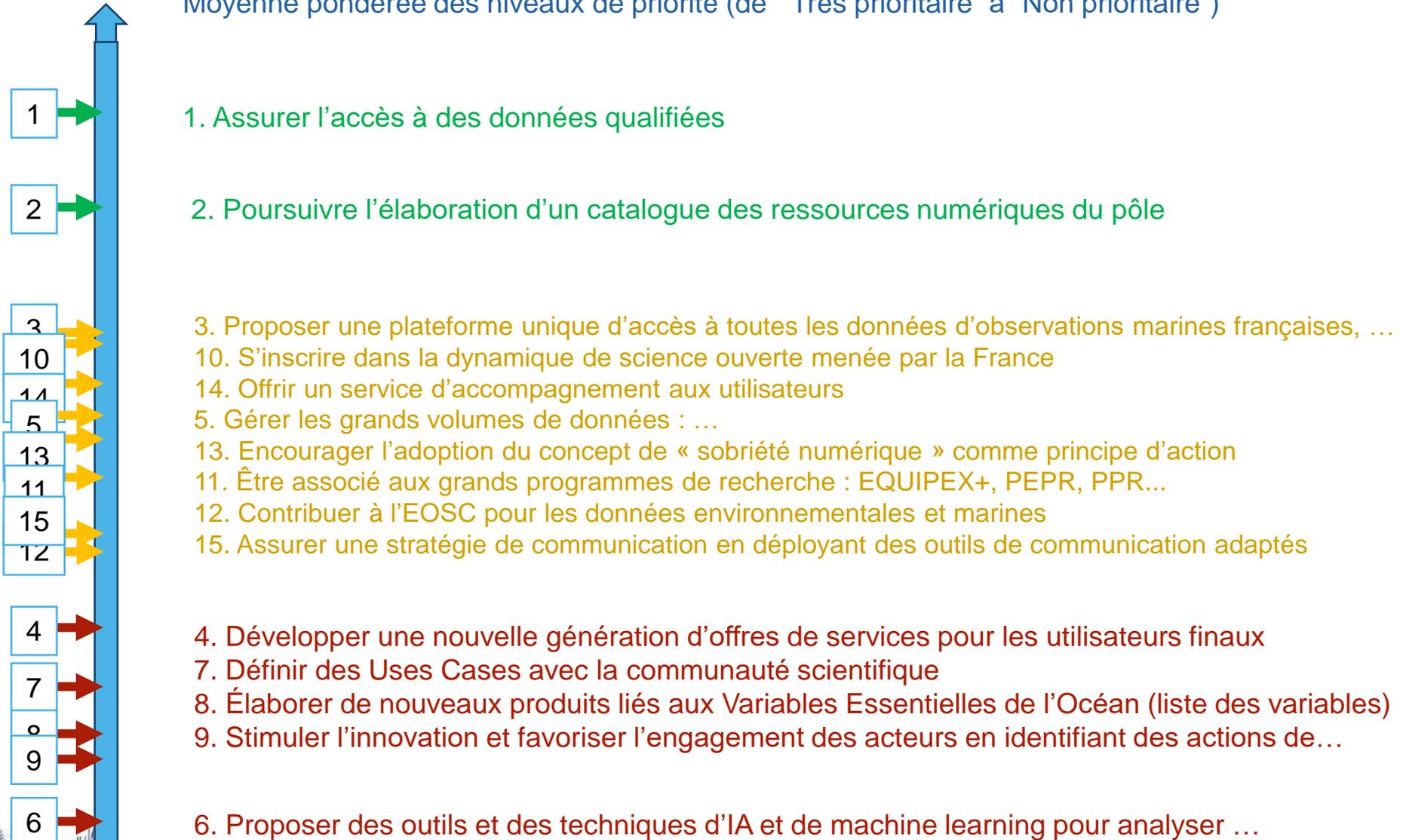
- Questionnaire pour construire collectivement la feuille de route ODATIS
- Profiter de l'AG ODATIS pour faire remonter les besoins des communautés
- 40 participants (représentants CODIR, CS, BE, Irs, programmes...)

A quelles communautés d'utilisateurs le pôle ODATIS doit-il s'adresser en priorité?



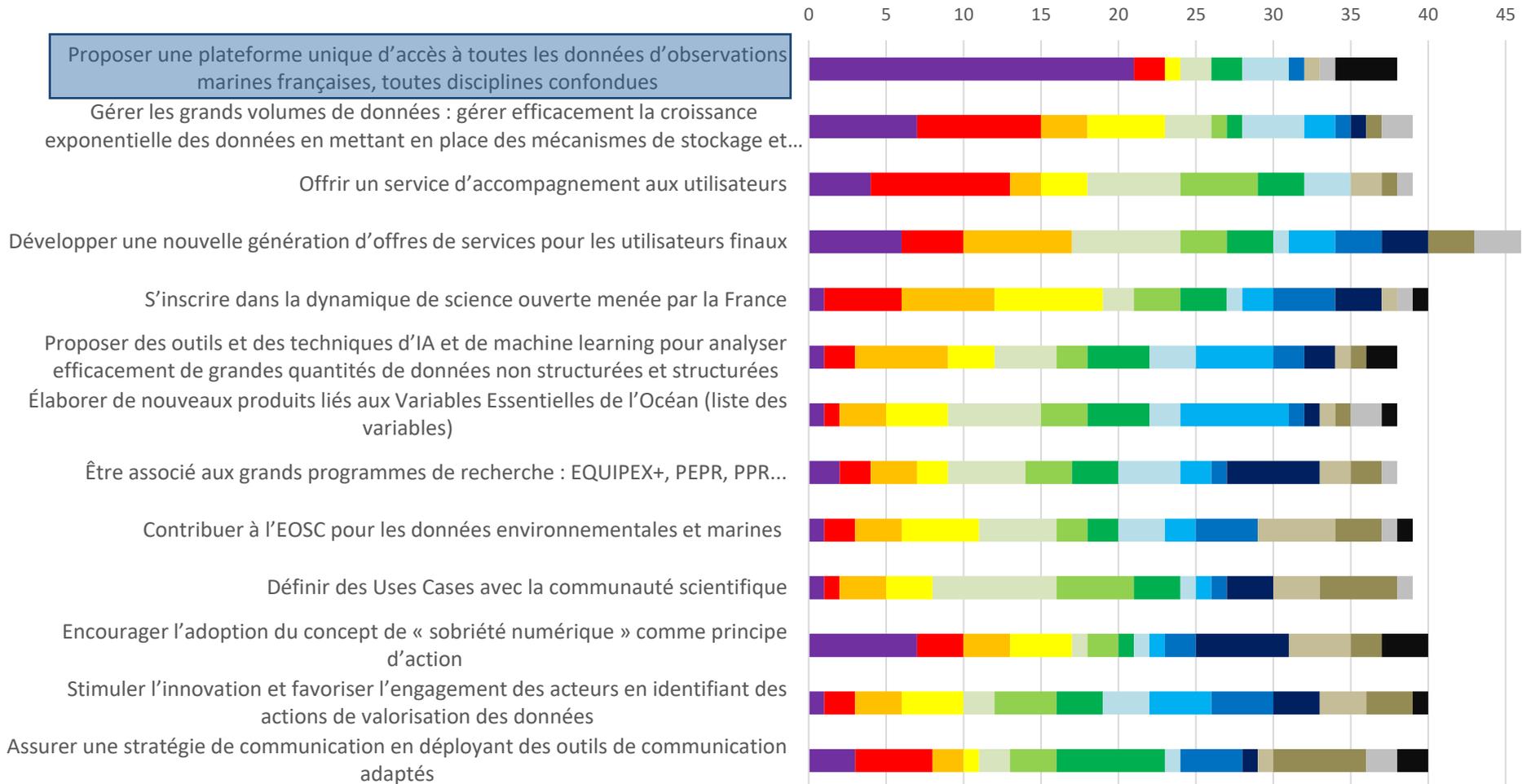
Les défis que le pôle ODATIS doit relever

Moyenne pondérée des niveaux de priorité (de “ Très prioritaire” à “Non prioritaire”)



Les défis que le pôle ODATIS doit relever

Pouvez-vous hiérarchiser ces 15 défis, par niveau d'importance?



■ 1 ■ 2 ■ 3 ■ 4 ■ 5 ■ 6 ■ 7 ■ 8 ■ 9 ■ 10 ■ 11 ■ 12 ■ 13 ■ 14 ■ 15

Poursuivre l'élaboration d'un catalogue des ressources numériques du pôle

Catalogue des données d'observation de l'océan et des produits dérivés, comprenant les données des IRs, des SNOs, des membres du pôle, des projets scientifique (PEPR, PPR...), catalogue des services, catalogue des outils et codes/scripts utiles à la communauté, ...

« Intégrer les **données d'expérience de laboratoire** en lien avec les sciences marines (Expérience en condition contrôlé de croissance du plancton par exemple) »

« s'assurer d'un **accès direct et immédiat aux données** à partir des moyens et services de traitement donner à la communauté les outils et moyens de diffusion et partage des codes/outils/tutoriaux (autogestion par les utilisateurs et non ODATIS) »

« Participer au partage d'un **modèle de métadonnées international**. manque dans les modèles existants (ISO-19115-3) pour pouvoir décrire précisément tous les types de données »

« **Catalogue des centres de calcul** »

« Clairement **montrer les liens** qui existent **entre les différents pôles** aux échelles nationale et internationale »

Développer une nouvelle génération d'offres de services pour les utilisateurs finaux

Outils d'analyses, outils de visualisations, outils de traitement des données, production d'indicateurs, ...

« **Personnes contact** pour les différents types de bases de données et de métadonnées (imagerie, chimie, biologie, climatologie...) »

« **préparation et transformation des jeux de données** (par ex pour construire des datasets d'apprentissage multivariés) »

« Outils d'estimation et de propagation des **incertitudes** dans les données »

« formations techniques »

« adopter ce qui sera mis en place aux niveaux des **ERIC** et **IR** océan (KPI, visualisation des séries via les ERDDAP...) »

« Outils de **qualification** des données »

« mettre en œuvre des **outils de sélection et d'extraction** ciblée sur une zone géographique à traiter »

« S'assurer du lien avec les **Jumeaux Numériques** à venir, Océan et côtier **Outils de traitement** des données hybrides (in situ, satellites, modélisation) »

« S'assurer de **l'interface avec les services européens** type Copernicus/CMEMS, CAMS etc. qui proposent le même type de produits/services. »

« Développer un **portail de recherche de données** basé sur le principe du portail isidore développé par la communauté shs (huma-num). <https://isidore.science/> »

Gérer les grands volumes de données

Données de drone, données d'imagerie, données vidéos, données de fibres optiques, données de modélisation, données de bio-informatique, ...

« Plutôt accéder aux données sans les stocker une autre fois (**pas de duplication**) »

« Veiller à ce que ces "données" (pour la modélisation je parlerais de résultats numériques!) ne soient pas déjà disponibles par ailleurs pour **limiter justement des duplications et leur stockage (bilan carbone)** »

« La gestion de données "volumineuses" passent par la définition des technologies de stockage associées.
Quelle technologie avons nous besoin pour les pôles de données (et pas que odatis)? **s3, webdav, irods, egi?**
Faut il disposer d'un seul grand centre pour stocker les données?
Il faudrait quantifier la volumétrie des données actuelles et leur taux de croissance pour vir en quoi les données sont volumineuses. »

« Il y a aussi tout le catalogue CMEMS, AVISO+, CORIOLIS.. mais justement, les données sont déjà organisées.
Je ne crois pas que revient à ODATIS de se doter (et doubler!) les immenses ressources de stockage et calcul pour gérer la croissance exponentielle des données. »

« **Données satellites** »

« données des systèmes de mesures "en route" de type **Ferry Box** »

« **Données gliders ou planeurs sous-marins, AUVs et flotteurs-profiteurs** »

« Quel sera le lien avec les **données génomique** (barcode, metabarcoding) ?

Il existe un portail OBA mis en place suite au projet TARA (<https://oba.mio.osupytheas.fr/ocean-atlas/>).

Est ce type de données qui devient de plus présent dans les réseaux d'observations (ex. MOOSE) et PPR FUTURE-OBS (observatoires augmentés) sera pris en compte par ODATIS ?

Comment assurer le lien entre les données océaniques et biologiques ? »

« **Données génétiques** (métabarcoding, métagénomique) »

Offrir un service d'accompagnement aux utilisateurs

Ateliers techniques, guides de bonnes pratiques, webinaires, séminaires dans des labos/OSUs, contribuer à l'offre de formations (ANF spécifiques océan), tutoriels en ligne, ...

« Mise en ligne de **courte vidéo** (max 10 minutes) une fois par semaine ou par mois pour expliquer un des aspects du pole (tuto, presentation des personnes qui travaille pour le pole, bilan des données du catalogue, ...) »

« Services permettant aux utilisateurs de construire et **partager ces ressources** pour des thématiques ou communautés **qu'ils définiront eux-mêmes** »

« Partage de **notebooks** »

« Je pense qu'il faut des **ateliers techniques dans les labos**. Une approche apprendre / découvrir "en faisant". »

« proposer une **graduation dans la proposition de services d'accompagnement**, du plus basique (standardisation, stockage des données) au plus élaboré (ex. traitement/hybridation de données). »

« Clarifier le "**chemin des données**", de l'envoi à la mise à disposition »

« **Qualification** et/ou validation des données, produits et services distribués »

« Formation spécifique pour les enseignants »

« **forum de questions** »

« faire **le lien avec OBPS** où sont recensés tous les BP ocean

(<https://www.oceanbestpractices.org>) »

Connaissez-vous des communautés dont les besoins ne sont pas couverts par les CDS d'ODATIS ?

« *couleur de l'eau* »

« *GIPREB* »

« *Oreme* »

« Communauté *paléo.* »

« *IUEM* »

« Quid du *MIO/OSU PYTHEAS* pour les données génomiques (ex. OBA) »

« Les écologues qui travaillent sur les communautés de microorganismes. Certains d'entre eux possèdent des données de biodiversité obtenues de stations d'observation dont les données hydrologiques sont dans les CDS ODATIS. Ce serait peut-être logique que leurs *données de biodiversité* (basées sur l'analyse de données génétiques) soient aussi stockées dans le CDS ODATIS. »

Un CDS qui a une expertise spécifique doit-il l'ouvrir à la communauté nationale ?

33 => oui

4 => non

Quelle est votre vision pour le pôle au cours des 5 prochaines années ?

« déjà remplir les objectifs identifiés ! » « **Consolider les acquis.** »

« **Conforter ses fonctions et ses missions** »

« **Offre de produits et de services plus claire** donc associer à un gros effort de communication voire de lobbying : Mise en avant du rôle de support incontournable pour la communauté scientifiques et les autres utilisateurs »

« Fidéliser les utilisateurs et en générer de nouveaux via la **qualité et l'utilité des produits distribués** »

« **Renforcer sa lisibilité** à l'échelle nationale. »

« implication de l'équipe dans des projets thématiques comme les **PEPR** »

« Travailler sur la **sobriété** »

« **diffuser les bonnes pratiques** dans la collecte, l'archivage et la validation des données. »

« **S'interfacer avec les schémas stratégiques du numérique (SSN) des OSU** »

« De continuer à promouvoir le traitement et l'archivage des données qualifiées, de mieux intégrer les différents type des données et **rendre le pôle incontournable également au niveau des IR/ERIC** »

« Il me paraît **illusoire de vouloir créer un mega-centre de données** qui stockerait l'ensemble des données sur le milieu marins, c'est un défi énorme et il faudrait faire table rase de toutes les autres sources de données IR, SNO, qui ont déjà mis en place de moyens de diffusion de leur données. »

« Je ne voit pas ODATIS comme un centre d'analyse et de stockage de données, ce qui demanderait des ressources énormes pour le faire bien, et qui doublerait des efforces déjà en place. »

Quels objectifs stratégiques pensez-vous que le pôle ODATIS devrait poursuivre au cours des 5 prochaines années ?

« Être **plus visible au niveau des IR** et des actions françaises dans les ERIC existants ou en construction »

« Assoir la **visibilité** et l'importance des activités du pôle auprès du **monde académique** et des **instituts** »

« Favoriser la visibilité des données via infrastructures déjà existantes, **SeadataNet, Emodnet** »

« Tout utilisateur doit pouvoir **facilement trouver les données disponibles** pour une période donnée, une zone géographique donnée et un paramètre donné, quelque soit la source de ces données. »

« Une **plateforme web claire** pour bien répertorier les données. »

« Faire davantage **le lien** avec les **données biologiques** (ex. EMBRC) afin que les utilisateurs puissent intégrer les données océanographiques et génomique par exemple » »

« assurer l'intégration avec les **services et plateformes européennes**, en particulier les clouds commerciaux (WEkEO, Creodias, ...) »

« **Formation & communication** sur ces données et outils »

« **Interaction avec des autres poles** - pour les données dans les frontières air-mer, océan- continent »

Pouvez-vous identifier des domaines spécifiques où le pôle devrait se concentrer davantage ?

« **Qualité des eaux, impacts anthropiques** »

« *Biodiversity, Human Activity* »

« **Accompagner et communiquer** »

« Renforcer le **soutien en RH** pour la valorisation /formattage de données ODATIS pour l'interopérabilité »

« Proposer une **offre de service claire et pérenne** pour l'ensemble des acteurs »

« Quel **levier ODATIS** pourrait avoir lors des demandes des **campagnes en mer** par exemple ? »

« Les interactions avec les autres centres de données / initiatives nationales / internationales concernant les **données de biodiversité.** »

« il serait essentiel de davantage motiver/influencer/forcer la communauté scientifique sur le **besoin d'archiver les données existantes.** »

Quels types de collaborations et de partenariats le pôle ODATIS devrait-il développer ?

« Centres de données nationales, européen, internationale »

« Avec la montée des **nouvelles techniques numériques** (IA, ML, hybridation) et **de la RSE** (sobriété), il serait intéressant de construire des partenariats stratégiques avec ces **nouveaux acteurs** en y associant les autres pôles de données via l'IR. »

« **Emodnet, EuroSea, Eoos** »

« **COPERNICUS, MERCATOR** »

« échanges avec des **partenaires européens** pour les bonnes pratiques »

« Les **collaborations avec les centres de données internationales** doivent être renforcées pour faciliter le transfert des données vers les bases de données internationales. »

« **Collaboration avec le PNDB** pour accompagner les biologistes marins dans la diffusion de leurs données. »

« visibilité auprès des **grandes agences européennes** (Copernicus, ESA, EUMETSAT, ECMWF,...) »

Comment jugez-vous l'implication du pôle ODATIS dans l'établissement de collaborations internationales ?

« **Insuffisant** mais en progrès »

« **faible** actuellement »

« Insuffisante et **peu claire** »

« il reste certainement des **marges de progression** pour faciliter le transfert de données vers des bases internationales »

« D'après mon expérience, **faible**...sans doute manque de communication, d'échanges entre les coordinateurs IR et ODATIS (ex. JERICO) »

« Je pense que même si ce genre d'implication peut s'avérer très intéressante, elle doit rester limitée afin de permettre à l'équipe ODATIS, qui n'est pas extensible, de se focaliser sur les priorités du pôle. »

« Succès de la **visibilité du portail AVISO au niveau international** »

« On voit plus ODATIS sur la **coordination nationale**. Peut-être est-ce aux organismes de tutelles d'établir les collaborations internationales. »

« Parfait pour la **mise en commun de méthodes communes** de travail au sein de la communauté Française »

Selon vous, quels ajustements seraient-ils nécessaires dans la structure de gouvernance ou l'organisation du pôle ODATIS pour atteindre les objectifs à long terme ?

« **Disponibilité RH indispensable et pérenne** »

« **Plus de RH dédié** »

« Il faudrait un **apport en moyens humains** à la hauteur des ambitions du pôle de la part des différents partenaires. »

« Un CS qui se reunit plus souvent s'il est prévu de **mettre en place un AAP** »

« Sans doute les **coordinateurs d'IR** devraient être impliqués dans le CS ou COPIL ou autre comité d'ODATIS »

Seriez-vous favorable à la mise en place d'un appel à projet de ce type pour le pôle ODATIS ?

- Les autres pôles de données de l'IR Data Terra proposent un appel à projet ouvert en continu afin de fournir à la communauté scientifique des moyens humains et matériels mutualisés pour répondre au mieux aux besoins autour de la gestion des données du Système Terre

22 oui – 3 non