



Aperçu des services fournis aux producteurs et utilisateurs de données marines par ODATIS, le pôle français de données océaniques

Sabine SCHMIDT ¹, Erwann QUIMBERT ², Caroline MERCIER ³,
Jean-François PIOLLE ², Mark HOEBEKE ⁴, Clémence COTTEN ²,
Dimitry KHVOROSTYANOV ⁵, Joël SUDRE ⁶, Cyril GERMINEAUD ⁷,
Dominique OBATON ², Frédéric MERCEUR ²

1. Univ. Bordeaux, CNRS, Bordeaux INP, UMR 5805, EPOC, F-33600 Pessac, France.
sabine.schmidt@u-bordeaux.fr
2. Ifremer, Centre de Bretagne, ZI Point du diable 29280 Plouzané, France.
Erwann.quimbert@ifremer.fr
3. Akkodis, 7 Boulevard Henri Ziegler, 31700 Blagnac, France.
4. Station Biologique de Roscoff, CNRS-Sorbonne Université, ABiMS FR2424, Place Georges Teissier, 29680 Roscoff, France.
mark.hoebeke@cnsr.fr
5. LOCEAN/IPSL, Sorbonne Université-CNRS-IRD-MNHN, 75005 Paris, France
6. UAR2013 CPST - IR DATA TERRA, Z.P. de Brégaillon, CS 20330, 83507 La Seyne sur Mer cedex, France.
joel.sudre@data-terra.org
7. CNES, 18 avenue Edouard Belin, 31400 Toulouse, France.
Cyril.Germineaud@cnes.fr

Résumé :

Le nombre et la variété des techniques d'acquisition de données nécessitent des méthodes efficaces pour améliorer la disponibilité des données bien décrites via des portails interopérables, qui facilitent leur partage selon les principes FAIR (Findable, Accessible, Interoperable, Reusable) pour les producteurs et les utilisateurs. ODATIS, le pôle de données et de services pour l'océan de Data Terra, l'infrastructure de recherche française pour les données du Système Terre et Environnement, constitue le point d'entrée pour accéder à toutes les données d'observation marine et littorale françaises. Après une présentation d'ODATIS, l'offre de service à destination des producteurs et des utilisateurs des données marines est décrite. Elle porte sur les différentes étapes du cycle de vie de la donnée depuis la planification de sa collecte ou de son acquisition, jusqu'à sa mise à disposition dans des environnements de recherche virtuels.

Mots-clés : Données marines, Cycle de vie des données, Plan de gestion des données, FAIR, et services autour des données, Interopérabilité, Normes, Standards.

Thème 3 – Instrumentation, mesures, imagerie et télédétection

1. Introduction

Les environnements côtiers, estuariens et littoraux sont soumis à de multiples pressions. D'une part, ils sont en première ligne pour faire face aux nombreuses conséquences du changement global sur l'océan : augmentation du niveau marin et de la température, tempêtes plus violentes, risque de submersion marine, érosion du littoral (CHUST *et al.*, 2010). D'autre part, ils sont très prisés : au moins 30 % de la population mondiale vit à moins de 100 kilomètres des côtes (NEUMANN *et al.*, 2015).

De nombreuses activités économiques y sont aussi concentrées : aux secteurs traditionnels tels que le transport maritime, le tourisme, la pêche, l'aquaculture et l'exploration d'hydrocarbures, s'ajoute le développement des énergies renouvelables à partir des forces de la mer. Tous ces facteurs sont susceptibles d'impacter l'hydrodynamique et la dynamique sédimentaire, la morphologie des fonds, la qualité des eaux et le fonctionnement des écosystèmes.

Une vision combinée des conséquences de ces multiples activités et pressions sur les régions côtières et littorales est difficile à établir en raison de la diversité et de la variabilité des environnements marins (<https://www.futureearthcoasts.org/>). Bien que le nombre d'observations marines et côtières ait nettement augmenté, que ce soit par des mesures *in situ*, en laboratoire ou par télédétection, chaque donnée reste à la fois coûteuse à acquérir et unique. Afin de tirer le meilleur parti de ces flux de données au profit de la connaissance et de la société, il faut développer des méthodes efficaces pour améliorer leur disponibilité et l'extraction des informations (SCHMIDT *et al.*, 2020).

Le nombre et la variété des techniques d'acquisition de données nécessitent le développement de portails interopérables, qui facilitent le partage des données selon les principes FAIR (*Findable, Accessible, Interoperable, Reusable*) pour les producteurs et les utilisateurs. Les principes FAIR visent à rendre les données *facilement trouvables* (ou découvrables), accessibles, interopérables et réutilisables dans une démarche de *Science Ouverte*.

Au niveau national, ODATIS, le pôle de données et de services pour l'océan de DATA TERRA, e-Infrastructure de Recherche dédiée au Système Terre et à l'Environnement, est le point d'entrée pour accéder à toutes les données d'observation marines françaises (Ocean Data Information and Services ; www.odatis-ocean.fr/en/). Après une présentation du pôle de données océan, l'objectif de cette communication est de porter à connaissance des communautés l'offre de service d'ODATIS.

2. ODATIS : le pôle de données marines

2.1 ODATIS dans le paysage national et européen

Créée en 2016, l'e-Infrastructure de Recherche DATA TERRA (<https://www.data-terra.org/>) est fondée sur quatre pôles de données et services correspondant à chacun des grands compartiments du système Terre dont ODATIS, auquel se rajoutent des services

transversaux comme DINAMIS, dédié à l'accès aux images satellites à haute résolution, et depuis 2024 le Pôle National de Données de Biodiversité. L'IR figure sur la feuille de route des infrastructures de recherche du Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation.

ODATIS est le point d'entrée pour accéder à l'ensemble des données françaises d'observation de l'océan. Ses objectifs sont de promouvoir et de faciliter l'utilisation des observations océaniques, de contribuer à la compréhension de l'océan et de ses interactions avec les autres composantes de la Terre. Les données gérées par ODATIS comprennent les variables de toutes les disciplines marines (physique, chimie, biologie...), quelle que soit la technique utilisée (satellites, observatoires *in situ*, campagnes de terrain, analyses en laboratoire). ODATIS a donc la charge de la gestion des données et métadonnées associées aux observations marines, du littoral au large, et de la surface au fond de l'océan, en physique, chimie et biologie. Les interfaces avec les autres domaines sont nombreuses, terre/mer sur le littoral, océan/atmosphère, océan/sous-sol océanique, et des moyens d'observation souvent mis en commun. Les tutelles du pôle sont au nombre de 8 : CNES, CNRS, Ifremer, IRD, SHOM, le réseau des Universités Marines, CEREMA et Météo-France.

La mission principale du Pôle Océan ODATIS est de mettre à disposition des données, des produits, des logiciels, des outils et des services destinés principalement aux communautés scientifiques et acteurs publics français travaillant dans le domaine de la recherche océanographique. Le pôle se doit de faciliter l'accès aux données satellitaires, aéroportées et *in situ* acquises et gérées, par les laboratoires de recherche ou les structures fédératives (OSU, FR,...), par des infrastructures nationales telles que les Services Nationaux d'Observation (SNO), les SOERE, la flotte océanographique, les avions et par les missions spatiales. Le pôle s'inscrit dans les paysages national et européen en articulation étroite avec les infrastructures en place et en construction. Il est impliqué dans des projets nationaux et européens comme porteur (ANR Flash Copilote) ou partenaire (PIA3 Gaia Data, Blue Cloud, ...)

2.2 Feuille de route du pôle ODATIS

Afin de favoriser l'utilisation des observations océaniques, la feuille de route d'ODATIS définit les actions suivantes :

- assurer la conservation à long terme de l'ensembles de données ;
- offrir une vision globale des observations *in situ* et satellitaires et de leurs produits dérivés ;
- faciliter l'accès, via un portail unique, à des bases de données entièrement décrites, localisées et qualifiées, en accord avec les normes scientifiques actuelles ;
- s'affranchir des frontières de l'espace, du temps et des disciplines en assurant l'interopérabilité des ensembles de données ;

Thème 3 – Instrumentation, mesures, imagerie et télédétection

- favoriser l'utilisation combinée de données de nature (*in situ*/satellite) ou d'origine (réseaux opérationnels/expériences scientifiques) différentes ;
- aider à extraire des informations des bases de données en proposant des outils d'exploration, d'extraction et d'analyse, et des environnements de recherche virtuels.

3. L'organisation du pôle de données Océan ODATIS

Pour mener à bien ses missions, ODATIS doit développer un dispositif global et performant d'accès et de traitement de données, produits et services dédiés à l'océan. ODATIS repose sur un ensemble de piliers décrit ci-dessous.

3.1 Le portail de données

Le portail d'ODATIS donne accès au catalogue. Ce catalogue propose plusieurs modalités d'accès aux données : recherche avec des filtres de sélection, description de la donnée, visualisation, et téléchargement (direct ou via le portail local des partenaires). Les données gérées étant par nature géoréférencées, les dispositions et protocoles relatifs à ce type d'informations géographiques s'appliquent : directive Inspire, métadonnées en norme ISO 19115, protocoles d'interopérabilité OGC...

Le catalogue du pôle de données Océan utilise les services et les outils de l'Infrastructure de Données Géographiques (IDG) Sextant. Le catalogue de données et de produits donne accès à des centaines de fiches de métadonnées pour accéder aux jeux de données fournis par la communauté nationale et à leurs liens avec les programmes internationaux.

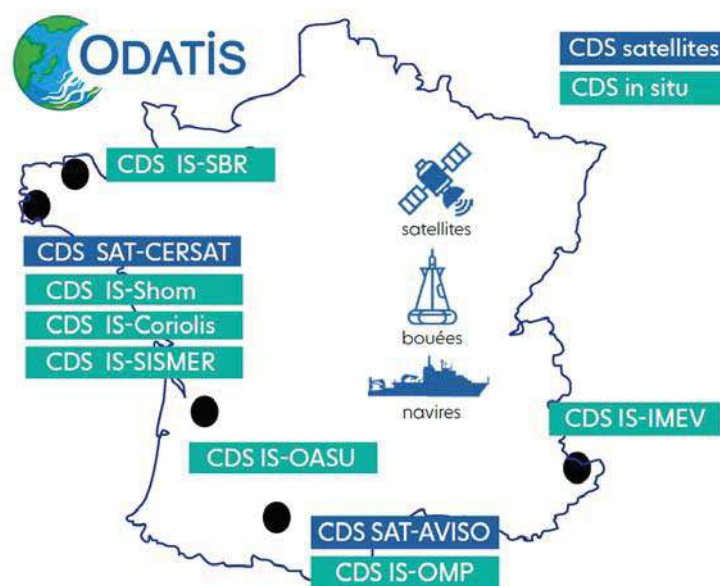


Figure 1. Les centres de données et service d'ODATIS.

3.2 Les Centres de Données et Services d'ODATIS

Pour atteindre ses objectifs, ODATIS s'appuie sur un réseau distribué de Centres de Données et de Services (CDS) opérés par les organismes de recherche partenaires du pôle (Figure 1). Actuellement au nombre de 9, les CDS assurent le travail de gestion et/ou traitement des données.

Les CDS *in situ* reposent sur des observations collectées avec des systèmes déployés depuis des navires ou engins dérivants, des ballons, des équipements fixes à terre ou en mer, ou des mesures en laboratoire. Ces données peuvent être produites par le biais de services d'observation, de chantiers ou de campagnes à la mer. Les CDS satellites reposent sur des observations collectées par des instruments et des capteurs installés sur des satellites.

Les CDS ODATIS gèrent et fournissent des jeux de données multi-capteurs, sur de longues séries temporelles homogènes, en appliquant les principes FAIR et en offrant la garantie de services sur la gestion de données (référencement, hébergement, diffusion, l'interopérabilité). Les CDS offrent la garantie de services sur la gestion des données : enregistrement des données, traitement des données, contrôle de la qualité des données, archivage à long terme et accès de routine aux données marines.

3.3 Les ateliers techniques

Le pôle de données Océan ODATIS propose chaque année plusieurs ateliers techniques (www.odatis-ocean.fr/activites/ateliers-techniques) réunissant les partenaires autour de présentations et tables rondes pour définir les orientations techniques du pôle. Des mises en pratique sont également organisées sur la base d'exercices d'application de plusieurs logiciels de visualisation, d'analyse et de traitement de la donnée. Les ateliers techniques sont annoncés sur le portail d'ODATIS et sont ouverts à tous.

3.4 Les Consortiums d'Expertise Scientifique

Enfin, ODATIS s'appuie sur des Consortiums d'Expertise Scientifique (CES) <https://www.odatis-ocean.fr/activites/consortium-dexpertise-scientifique>.

Les CES sont des regroupements d'experts menant des travaux de recherche ou développant des méthodes innovantes de mobilisation des données d'observation, et capables d'assurer la réalisation de prototypes de produits à valeur ajoutée, ou la mise en œuvre de ces prototypes pour produire des données spécifiques sur des problématiques Océan (littorales, côtières ou hauturières), autour de champs thématiques (processus physiques, chimiques et biologiques, écosystèmes, échanges océan/atmosphère, approches globales, exploitation raisonnée des ressources...) pour répondre aux défis sociétaux et environnementaux de notre époque. Actuellement au nombre de 5 (pH/CO₂ marin, Oxygène dissous, Couleur de l'océan, Cytométrie en flux et Imagerie optique benthique), les CES sont créés à la demande des communautés scientifiques.

4. L'offre de service d'ODATIS

L'offre de service d'ODATIS porte sur les différentes étapes du cycle de vie de la donnée. Le cycle de vie de la donnée désigne les étapes et les flux par lesquels une donnée passe depuis la planification d'un projet et sa collecte jusqu'à sa publication, voire sa réutilisation (Figure 2). Les données ont souvent une durée de vie plus longue que le projet de recherche qui les a créées. De nombreux appels d'offre nationaux et européens exigent désormais que les projets lauréats produisent un plan de gestion des données (PGD, *Data Management Plan DMP*).



Figure 2. Le cycle de vie de la donnée.

Lors du montage d'un projet, les porteurs peuvent prendre contact avec ODATIS pour envisager les conditions d'assistance dont ils pourraient bénéficier en matière d'élaboration de plan de gestion de données (PDG), de bancarisation et d'analyses de données environnementales (Figure 3). Un aspect important concerne la bonne description des données qui doit comprendre des métadonnées riches et un vocabulaire standardisé, respectant les standards internationaux. Un premier enjeu est de mettre en œuvre les principes FAIR, à rendre les données plus visibles et accessibles en étant référencées dans le catalogue ODATIS, et mieux suivies et citées grâce à un Identifiant d'Objet Numérique (DOI).

ODATIS propose également un service de publication de données scientifiques ouvertes sur la mer, par l'intermédiaire de SEANOE (www.seanoe.org), qui fournit un DOI. Ce DOI correspond à un identifiant numérique propre au jeu de données, rattaché à une URL pérenne et dans lequel l'auteur (le producteur de la donnée) est identifié clairement. Cela

permet une citation précise, fiable et pérenne du jeu de donnée, ainsi la donnée est plus facilement (re)trouvable et réutilisable (Findable et Reusable du principe FAIR).

ODATIS dispose d'un ensemble d'outils et de logiciels et d'Environnements de Recherche Virtuels (VRE) mis à disposition des producteurs de données et des utilisateurs afin de traiter, comparer, extrapoler, visualiser et valider un ensemble de données afin que ces dernières puissent être valorisées dans des publications de recherche et des rapports ou être introduites dans des modèles (visualisation à long terme, prévision climatique, etc.).

Pour tirer le meilleur bénéfice des flux de données croissants, ODATIS propose des moyens pour analyser et traiter de grands volumes de données à distance sur les centres de calcul à haute performance (HPC). Des accès aux clusters de calcul du CNES, HAL, et de l'Ifremer, DATARMOR, sont possibles sur la base d'une description détaillée du besoin en lien avec ODATIS (<https://www.odatis-ocean.fr/contact>).



Figure 3. Du producteur à l'utilisateur des données.

5. Conclusion

Depuis la révolution industrielle, l'empreinte des activités humaines sur l'environnement mondial s'est accrue. Les conséquences actuelles et attendues du changement global sur l'océan sont multiples. Des observations sont nécessaires à tous les stades du processus scientifique : description, compréhension, modélisation et prévision. L'acquisition des données marines est souvent difficile et onéreuse. Elle nécessite d'accéder à des sites distants et d'utiliser de nombreux moyens techniques (flotte océanographique, sites instrumentés, *gliders*, bouées Argo...). Or, sans archivage, plus de 30 % des données sont perdues ou inutilisables 10 ans après leur acquisition. La préservation et la réutilisation des données marines constituent donc un enjeu majeur. Les scientifiques doivent relever ce défi dans un esprit d'éthique, de transparence et de reproductibilité des résultats. Dans ce contexte, l'application des principes FAIR et de la Science Ouverte aux données d'observation de la Terre est une priorité absolue pour ODATIS.

ODATIS, le pôle de données et de services pour l'océan de l'IR DATA TERRA, mène une démarche de communication afin d'informer les communautés scientifiques et

Thème 3 – Instrumentation, mesures, imagerie et télédétection

techniques autour de son offre de service autour des données. Les besoins des politiques publiques sont également pris en compte. Ainsi cette communication vise à donner un bref aperçu des outils existants et à favoriser les interactions pour poursuivre le développement du portail. Le portail d'ODATIS est riche d'informations et les communautés sont fortement incitées à le visiter pour y découvrir les services et les outils mis à leur disposition et le cas échéant à contacter le pôle. A titre d'exemple, dans le cadre de sa politique de science ouverte et du plan national pour la science ouverte, l'Agence Nationale de la Recherche exige désormais que tous les projets financés produisent un plan de gestion des données (PGD) qui est considéré comme un livrable. Les porteurs sont incités lors de l'élaboration de leur projet à prendre contact avec ODATIS, voire l'IR Data Terra, pour envisager les conditions d'assistance dont ils pourraient bénéficier en matière d'élaboration de PDG, de bancarisation et d'analyses de données environnementales.

L'offre de service d'ODATIS concerne bien évidemment les spécialistes en dynamique du littoral, génie côtier, risques et ressources littorales, et gestion raisonnée du littoral. Sa première action est de porter à connaissance les jeux de données existants via son catalogue dans les zones littorales et côtières dont celles acquises par les SNOs de l'IR ILICO. Compte-tenu de ses tutelles, l'offre de service d'ODATIS s'adresse aux acteurs académiques mais est ouverte aux industriels. ODATIS peut aussi promouvoir le développement de produits spécifiques à la carte. Par exemple, le CES Couleur de l'Océan avait exprimé le besoin en produits satellitaires moyenne résolution (MR) qui sont maintenant accessibles pour tous via un geobrowser (<https://odatis.acri-st.fr/>). Le traitement de données de couleur de l'océan est réalisé par ACRI-ST sous contrat CNES en exploitant des algorithmes innovants appliqués à des données à moyenne résolution spatiale (environ 300m). Il est, pour l'instant, focalisé sur les eaux côtières de France métropolitaine.

6. Références bibliographiques

- NEUMANN B., VAFEIDIS A.T., ZIMMERMANN J., NICHOLLS R.J. (2015) *Future coastal population growth and exposure to sea-level rise and coastal flooding – a global assessment*. Plos One 10, e0118571. <https://doi.org/10.1371/journal.pone.0118571>
- CHUST G., CABALLERO A., MARCOS M., LIRIA P., HERNÁNDEZ C., BORJA A. (2010). *Regional scenarios of sea level rise and impacts on Basque (Bay of Biscay) coastal habitats, throughout the 21st century*. Estuar. Coast. Shelf. R., 87, 113-124, <https://doi.org/10.1016/j.ecss.2009.12.021>
- SCHMIDT S., MAUDIRE G., NYS C., SUDRE H., HARSCOAT V., DIBARBOURE G., HYUNH F. (2020) *Streamlining data and service centres for easier access to data and analytical services: the strategy of ODATIS as the gateway to French marine data*. Frontiers in Marine Science 7: 548126. <https://doi.org/10.3389/fmars.2020.548126>