

---

# **PLATEFORME DES RESEAUX DE LA MISSION POUR LES INITIATIVES TRANSVERSES ET L'INTERDISCIPLINARITE**

Synthèse

GT DONNEES

2020

Pad de rédaction GT-données v2.0 18/12/2019 pour rappel la cartographie des actions des réseaux : <http://www.cnrs.fr/mi/IMG/pdf/gtinterreseaux-cartosynthesev6.pdf>

Cette nouvelle version v2.0 s'apparente davantage à un centre de ressources éditorial inter-réseau autour de la gestion de données... dans laquelle on rassemble les productions de chaque réseau en matière de gestion de données [la 1ère version avec des idées clés, plus textuelle qu'on a abandonnée est là : <https://etherpad.in2p3.fr/p/gbp-donnees>]

Objectifs: rassembler l'ensemble des productions (formation, conférences, séminaires, présentations...) de chaque réseau apparentées avec la gestion des données, par phase du cycle de vie des données

Le think tank : - Olivier Brand-Foissac et Sophie Nicoud : Resinfo - Joanna Janik : DIST (Renatis) - Maurice Libes (SIST) - Violaine Louvet : Calcul - Caroline Martin : <- quel réseau? - Alain Rivet : QeR - Genevieve Romier : RBdD

[[Il me semble que l'on ne pourra pas se départir de recommandations générales même si celles-ci ont déjà été abordées par ailleurs, +1; recommandations adossées aux URL/Evenement organisés par les réseaux, ce qui en fait l'originalité]]

===== Question sur des options possibles quant à la finalité de ce guide : Fait-on de ce guide aussi un outil de "promotion/valorisation" des actions du réseau en plus d'un guide pratique ? Avec l'option de se dire : - Voici les recommandations que peuvent faire les réseaux professionnels sur la gestion et le partage des données partant des actions menées en leurs seins, des besoins professionnels éprouvés dans les pratiques par les communautés pour gérer les données et des échanges d'expériences inter-réseau ? Et si oui ajoute-t-on l'idée que cela pourrait être le reflet des activités présentes et passées des réseaux autour des données (idée de témoigner de l'évolution de nos actions et de la thématique dans le temps) ou - voici un guide purement "technique" sur la gestion des données avec l'objectif de faciliter la gestion et de donner accès à des ressources et point barre. [les 2 ne me semblent pas antinomiques.. c'est un peu l'orientation actuelle] =====

## 1. Imaginer

Contexte: Les réseaux participent à cette étape du cycle de vie en identifiant les problématiques techniques, en proposant des outils technologiques les plus pertinents pour accompagner la construction, la mise en œuvre du projet et la valorisation des résultats. Pour réaliser cet accompagnement, les réseaux organisent une veille sur les nouvelles technologies en émergence permettant le traitement et l'analyse des données, sur les nouveaux supports de valorisation de la recherche, les contraintes et opportunités qu'elles soient réglementaires, juridiques (politiques d'Open Data, H2020 par exemple) ou bien techniques, technologiques (TDM, ...) ou encore organisationnelles (plans de gestion de données). L'apport des réseaux dans cette étape est

également important en termes de croisement des disciplines et des métiers permettant un accompagnement de haute valeur

Faire de la veille technologique et réglementaire Assurer une veille réglementaire (entre autres données susceptibles d'avoir besoin de protection particulière (données de santé par ex.) directive Inspire Plan pour la science ouverte (voir) Assurer une veille technologique (standard de métadonnées, de formats et de protocoles d'échange , standards existants dans le domaine/discipline/métier, référentiels...) Quels méta données normes et standards pour les données (de calcul) et autres données Assurer une veille régulière sur les évolutions des besoins des utilisateurs scientifiques Assurer une veille par les réseaux métiers => un encart sur la MITI et liste des métiers veille RDA Assurer une veille sur les évolutions des infrastructures, en particulier sur ce qui se fait dans EOSC <https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud> Assurer une veille sur les services proposés aux utilisateurs scientifiques.

Activités de recherche et gestion des connaissances - Alain Rivet (CERMAV), Participer à l'organisation du management des données de la recherche : gestion de contenu et documentation des données - 6-8 juil. 2016 Paris (France) Vidéo - <https://youtu.be/UhBwjJDQcdg> [https://youtu.be/6laYJ\\_Rvm28](https://youtu.be/6laYJ_Rvm28) <https://youtu.be/7apGMS9gg5g> <https://youtu.be/Ld5vEBYAMNA> [https://youtu.be/wVgQ\\_2fM10s](https://youtu.be/wVgQ_2fM10s) ppt - [https://anfdonnees2016.sciencesconf.org/data/pages/donnees\\_renatis\\_A](https://anfdonnees2016.sciencesconf.org/data/pages/donnees_renatis_A)

Le défi organisationnel de la gestion des données dans les laboratoires et les établissements. - Eric Quinton (IRSTEA), Participer à l'organisation du management des données de la recherche : gestion de contenu et documentation des données - 6-8 juil. 2016 Paris (France) Vidéo : <https://youtu.be/vuUEv0MbYrM> <https://youtu.be/sHReS-G4Mg> <https://youtu.be/u9P5ODxeLmE> ppt [https://anfdonnees2016.sciencesconf.org/conference/anfdonnees2016/pages/dmp\\_analyse\\_securite.pdf](https://anfdonnees2016.sciencesconf.org/conference/anfdonnees2016/pages/dmp_analyse_securite.pdf)

Savoir anticiper les nouveaux outils, les technologies émergentes en matière (ANF 2016) <https://indico.mathrice.fr/event/5/contribution/3/material/slides/0.pdf>

\*Traçabilité des données Mettre en oeuvre les bonnes pratiques de traçabilité des activités de recherche : [http://qualite-en-recherche.cnrs.fr/IMG/pdf/guide\\_tracabilite\\_activites\\_recherche\\_gestion\\_conna](http://qualite-en-recherche.cnrs.fr/IMG/pdf/guide_tracabilite_activites_recherche_gestion_conna)

\*Traitement juridique des données

Résultats de la recherche et open data : le cadre juridique - Nathalie Gandon (INRA) Fredocs 2018 - Démarches innovantes en IST : expérimenter, proposer, (se) réinventer 3-5 octobre 2018, Albi Présentation de l'impact des Lois dites Valter (2015) et Lemaire (2016) sur les données de la recherche. pdf - [https://fredoc2018.sciencesconf.org/data/pages/Open\\_data\\_cadre\\_juridique\\_N\\_Gandon.pdf](https://fredoc2018.sciencesconf.org/data/pages/Open_data_cadre_juridique_N_Gandon.pdf)

Connaitre le paysage de la gestion des données Obtenir une vision d'ensemble de l'utilisation des données, au niveau du labo et de la discipline nécessité de se servir de catalogues internes ou externes, moissonnage des datasets connaitre les portails nationaux et Européens : Pôles de données (Odatis, Theia, Aeris, Form@ter ) de l'IR Data Terra connaître les licences relatives aux

jeux de données pour bien utiliser les données connaître les façons de citer les données

Identifier les parties des données susceptibles d'avoir besoin de protection particulière (données de santé par ex.) et la manière de les organiser (pseudonimisation ou anonymisation) afin de simplifier leur traitement par les scientifiques et d'éviter les identifications par croisement. suivi des appels Savoir identifier les données de simulation à conserver : peut-on reproduire le calcul ? Quel coût stockage vs calcul ? Lien avec la reproductibilité

- se former -liste de formation en fonction des métiers et des domaines du cycle de vie des données -les ANF montées par les réseaux métiers sont des vecteurs importants de l'état de l'art dans une discipline
- Capitalisation des supports de formation sous une forme ou une autre ? (accessibles sur les sites des réseaux, enregistrement des présentations pour visionnage ultérieur.)
- IDENTIFIER LES COMPETENCES et Expertises ( se rapprocher) Assurer des fonctions de mise en relation (BDD, équipes, technos, disciplines) -au sein des acteurs d'un projet : utilisation de méthodes et produits de gestion de projets
- intégrer des réseaux métiers, réseaux technologiques

Gestion des données et principes FAIR

FaiR ou Fair-RR ? Réutilisation vs Recherche Reproductible ? Atelier Dialogu'IST, 3 décembre 2019 Dialogu'IST vous propose dans la continuité un 8e atelier, plus particulièrement sur la Réutilisation des données dans le contexte de la Recherche Reproductible et de la Science Ouverte. Après une première partie rappelant le contexte et les enjeux, la seconde partie de la session présentera différents retours d'expériences de formations pour parvenir à la production de données réutilisables, répliquables et reproductibles. Des interventions orales, avec des témoignages et des retours d'expériences, seront proposées aux scientifiques – chercheurs, enseignants-chercheurs, doctorants, ingénieurs – ainsi qu'aux professionnels de l'information. Basé sur la dynamique d'échanges libres et spontanés, cet atelier-formation fera ainsi intervenir tous les acteurs concernés. Afin de donner la parole au plus grand nombre, cet atelier aura lieu en visioconférence, sur les sites organisateurs et sur des sites collaborateurs. Vous pouvez contacter les différents responsables de sites pour nous rejoindre. <http://renatis.cnrs.fr/spip.php?article520>

FAIRe vivre les données de recherche : retours d'expérience sur l'organisation, la préservation et le partage Atelier Dialogu'IST, 13 décembre 2018 Après une intervention de Francis André, coordinateur de l'alliance RDA France, sur le paysage politique et culturel de la Science Ouverte, seront exposés différents retours d'expériences sur la mise en œuvre de gestion des données et des principes FAIR, de DMP, etc. Professionnels de l'IST, Ingénieurs de Recherche scientifiques ou informaticiens présenteront leurs projets de plateforme de gestion des données et leur utilisation, les conséquences sur les comportements et métiers ainsi que des « trucs et astuces » pour la

réussite de tels projets. <http://renatis.cnrs.fr/spip.php?article519>

[Calcul]

[RBdD] MANDRAN, Nadine, 2016. Méthodes pour intégrer l'utilisateur dans la construction des applications. In : Système d'information embarqué, cahier/carnet de terrain et de laboratoire électronique : quelles interactions avec les bases de données ? [en ligne]. Paris. 10 mai 2016. [http://rbdd.cnrs.fr/IMG/pdf/utilisateurs\\_anfrbdd.pdf?171/5b534aea6e9bfe597fcf71ff840c5cf48c67033c](http://rbdd.cnrs.fr/IMG/pdf/utilisateurs_anfrbdd.pdf?171/5b534aea6e9bfe597fcf71ff840c5cf48c67033c)

MARTIN, Chloé, 2016. Informations juridiques liées aux données. In : Formation « Comment concevoir une base de données » (6-7 juin 2016) [en ligne]. Villejuif. 6 juin 2016. [http://rbdd.cnrs.fr/IMG/pdf/informations\\_juridiques\\_lies\\_aux\\_donnes.pdf?158/da55c34f313c136bdd51618f48](http://rbdd.cnrs.fr/IMG/pdf/informations_juridiques_lies_aux_donnes.pdf?158/da55c34f313c136bdd51618f48)

ANDRE, François, 2014. Normes OGC. In : ANF2014 DEVLOG-RBDD : Infrastructure de Données géographiques et Spatialisées [http://devlog.cnrs.fr/\\_media/ids2014\\_presentationogc\\_francoisandre.pdf?id=ids20](http://devlog.cnrs.fr/_media/ids2014_presentationogc_francoisandre.pdf?id=ids20)

BALY, Isabelle et GRISON, Philippe, 2014. Chaîne opératoire de réalisation d'une base de données. In : ANF CAIRN - rBDD 06/06/2014 Comment concevoir une base de données en archéométrie ? [http://rbdd.cnrs.fr/IMG/pdf/anfcairn-rbdd\\_5et6juin2014\\_peti.pdf?190/e455c772dd0f972d6b785fc9662cd8bfe](http://rbdd.cnrs.fr/IMG/pdf/anfcairn-rbdd_5et6juin2014_peti.pdf?190/e455c772dd0f972d6b785fc9662cd8bfe)

[QeR]

[Resinfo] Outils de reproductibilité (JoSy 2019) <https://indico.mathrice.fr/event/165/contribution/2/material/slides/0.pdf>  
<https://webcast.in2p3.fr/video/sidus>

[RENATIS] Mes commentaires en orange sont provisoires et à destination de Joanna pour prochaine discussion sur ce qu'il est important de faire figurer dans le guide ou de mentionner rapidement ... Et mes réponses en vert :)

Gestion des données et Open Science / contexte et enjeux pour resituer le partage des données [préciser qu'il est important de veiller sur les enjeux et de bien comprendre le contexte français, européen et international autour de la question des données ?]

Gestion des données de la recherche dans le contexte d'Open Science - Francis André (CNRS DIST) - Participer à l'organisation du management des données de la recherche : gestion de contenu et documentation des données - 6-8 juil. 2016 Paris (France) Science du 21ème siècle, science des données : dans un contexte d'omniprésence du numérique, et face aux défis sociétaux actuels, l'organisation d'une gestion raisonnée des données de recherche est un enjeu de taille pour les communautés scientifiques. L'Open science offre un cadre stratégique pour développer infrastructures et services de données conduisant à une meilleure circulation des résultats scientifiques. Petit tour d'horizon...

Vidéo - <https://anfdonnees2016.sciencesconf.org/resource/page/id/7> ppt - <https://anfdonnees2016.sciencesconf.org/resource/page/id/7>  
A enlever je pense

Gestion des données de la recherche dans le contexte d'Open Science - Francis André (DIST-CNRS)  
Participer à l'organisation du management des données de la recherche : gestion de contenu et

documentation des données - 3-6 juil. 2017 Paris (France) <https://anfdonnees2017.sciencesconf.org/data/pages>  
[plus récente mais pas de vidéo donc à voir ] OK

État des lieux / prospection

Big data et métiers de la documentation (J.M. Saläun, ENS Lyon) - Frédocs 2015 - Mieux accompagner la recherche : réalités d'aujourd'hui et perspectives pour les fonctions IST - 29 septembre - 2 octobre 2015, Saintes-Foy-Lès-Lyon Le mouvement dit des "big data" est souvent présenté comme une révolution susceptible de transformer radicalement des pans entiers de l'activité économique. A-t-il une conséquence sur celle des documentalistes ? Bien des caractéristiques de ce mouvement sont familières au monde de la documentation, à commencer par sa structure. Mais la documentarisation de tous les objets et de tous les événements issue de leur numérisation bouleverse les méthodes traditionnelles du métier par les effets d'échelle et les capacités de calcul et modifie sans doute aussi notre relation au savoir. Il est alors indispensable de renouveler les compétences pratiques et de se maintenir éveillé face au renouvellement de la demande de connaissances, même si le raisonnement documentaire reste pérenne. [https://fredoc2015.sciencesconf.org/conference/fredoc2015/pages/Salaun\\_FREDOC2015.pdf](https://fredoc2015.sciencesconf.org/conference/fredoc2015/pages/Salaun_FREDOC2015.pdf)

Politique d'accompagnement à la gestion des données

« Les politiques d'accompagnement des données : une comparaison internationale » : Simon Hodson (ISCU-CODATA) - Frédocs2013 - Gestion et valorisation des données de la recherche - 7 au 10 octobre 2013, Aussois [http://renatis.cnrs.fr/IMG/pdf/Hodson\\_FReDOC\\_Presentation\\_FR.pdf](http://renatis.cnrs.fr/IMG/pdf/Hodson_FReDOC_Presentation_FR.pdf) (Très complet - Panorama des différentes politiques - Tendances - Vision très prospective dans laquelle nous sommes projetés aujourd'hui - considérations sur les défis et les obstacles à lever pour la mise en place des données de la recherche vu par CO DATA un organe international) Intéressant parce que c'est S. Hodson ! Et parce que c'est lui qui a attiré (notre) attention à la nécessité d'avoir des actions du terrain (bottom) qui rejoignent celles de la gouvernance (up).

« From data management policy to implementation : opportunities and challenges for libraries » : Susan Reilly (Liber) intervention en anglais - - Frédocs2013 - Gestion et valorisation des données de la recherche - 7 au 10 octobre 2013, Aussois <http://renatis.cnrs.fr/IMG/pdf/ReillyFredoc.pdf>

Évolution des pratiques en IST pour accompagner le changement vers la gestion des données (mais pas que) - Vision compétences et formation [Peut-être en dire deux mots en amont des bonnes pratiques pour souligner l'importance de l'évolution nécessaire / la transition de nos métiers et l'évolution des compétences pour bien gérer les données et mieux accompagner les chercheurs - Vision évolution de la politique des établissements Pas de place dans un plan comme le notre pour ce guide selon moi. Manque la partie Accompagnement ? (pas la même chose que Formation)

Retour sur l'enquête Dialogu'IST : positionnement perçu et attendu des acteurs de l'IST face à l'innovation dans le contexte de la science ouverte en mouvement - Chris-

tine Hadrossek (RIATE), Sylvie Grésillaud (INIST-CNRS) - Fredocs 2018 - Démarches innovantes en IST : expérimenter, proposer, (se) réinventer 3-5 octobre 2018, Albi pdf - <https://fredoc2018.sciencesconf.org/resource/page/id/5>

Évolution et nouvelles pratiques autour du thésaurus PACTOLS de Frantiq pour son usage notamment en édition numérique - Blandine Nouvel (Centre Camille Jullian, Aix Marseille) - Fredocs 2018 - Démarches innovantes en IST : expérimenter, proposer, (se) réinventer 3-5 octobre 2018, Albi La réorganisation du thésaurus PACTOLS pour l'archéologie et ses capacités d'interopérabilité répondent aujourd'hui aux besoins de partage des données, au-delà de l'indexation des notices du Catalogue Collectif Indexé (CCI) de la Fédération et Ressources sur l'Antiquité (Frantiq) pour laquelle il a été conçu. Il s'ouvre à de nouveaux usages pour la gestion des données de la recherche et pour l'édition. On présentera le dispositif de transformation structurelle du thésaurus et le programme de sensibilisation et de formation qui l'accompagne. pdf - <https://fredoc2018.sciencesconf.org/resource/page/id/5> OK dans quelle partie ? (voir avec Stéphane ?)

Intervention sur la mutation des pratiques numériques et propositions de services à la MSH de Nantes Amélie Renard (MSH de Nantes) - Fredocs 2018 - Démarches innovantes en IST : expérimenter, proposer, (se) réinventer 3-5 octobre 2018, Albi La MSH Ange-Guépin mène depuis plusieurs années une politique volontariste dans l'accompagnement des équipes de recherche ligériennes à la transition numérique. Dans ce but, son pôle numérique s'est doté de deux plate-formes technologiques, Humanum-Loire et Progedo-Loire, chargées d'animer les échanges avec les chercheurs tout en leur proposant une offre variée de formations (courtes/longues, pratiques/théoriques, programmées/à la demande) visant à décomplexer leur rapport au numérique. Cette présentation sera l'occasion d'un retour d'expérience (leviers à actionner, contraintes, mesures de réussite...) ouvert aux discussions. pdf - <https://fredoc2018.sciencesconf.org/resource/page/id/5> même remarque que l'intervention précédente

Promesses et limites d'un séminaire interne à un laboratoire de SHS pour sensibiliser et former les chercheurs aux plans de gestion des données - Emmanuelle Morlock (HiSoMA, Lyon) - Fredocs 2018 - Démarches innovantes en IST : expérimenter, proposer, (se) réinventer 3-5 octobre 2018, Albi En matière de gestion et de partage des données de la recherche, certaines actions de sensibilisations peuvent être perçues comme des discours d'injonctions teintés de culpabilisation, qui au mieux n'ont aucun impact sur les pratiques, au pire ont pour effet de renforcer les réticences à l'opposé de ce qui était attendu. Or aucune politique de gestion, d'archivage et de partage ne pourra véritablement se mettre en place sans l'adhésion des chercheurs. Cette présentation décrira une expérience en cours au sein du laboratoire HiSoMA, un laboratoire de SHS spécialisé sur l'étude de la méditerranée antique pour impliquer chercheurs, doctorants et ingénieurs dans une démarche d'élaboration collective de plans de gestions de données, construite autour

d'un séminaire interne. pdf - <https://fredoc2018.sciencesconf.org/resource/page/id/5>

Réduire les distances et accompagner les mutations dans la production et la gestion des données de recherche : les enjeux de la transmission - Stéphane Loret et A. Renard (MSH Nantes) - Participer à l'organisation du management des données de la recherche : gestion de contenu et documentation des données - 3-6 juil. 2017 Paris (France) ppt - [https://anfdonnees2017.sciencesconf.org/data/pages/ANF\\_RENATIS.pdf](https://anfdonnees2017.sciencesconf.org/data/pages/ANF_RENATIS.pdf)

Du dépôt au partage des archives sonores de la recherche, les évolutions du métier de phonothécaire - Françoise Acquier (CRESSON, Ecole d'architecture Grenoble), Ariane Néroutidis, Marine Soubrié (Phonothèque Aix en Provence) - Frédocs 2015 - Mieux accompagner la recherche : réalités d'aujourd'hui et perspectives pour les fonctions IST - 29 septembre - 2 octobre 2015, Saintes-Foy-Lès-Lyon L'accès aux données de la recherche est aujourd'hui au coeur des métiers de l'IST au CNRS. Depuis la fin des années 1970 deux laboratoires de recherche se préoccupent de l'archivage des données enregistrées sur le terrain : la Phonothèque de la Maison Méditerranéenne des Sciences et de l'Homme à Aix-en-Provence et le Centre de Recherche sur l'Espace Sonore et l'Environnement Urbain à Grenoble. L'intervention a pour objectif de présenter comment le métier de phonothécaire se redéfinit autour de nouvelles pratiques : la contextualisation des archives, l'exigence des formats des données, leur interopérabilité, leur archivage à long terme, leur éditorialisation sur des Carnets de la plateforme Hypothèses, la mise en œuvre de guide de bonnes pratiques alliant éthique et droit. Les archives sonores participent ainsi pleinement de l'essor des "digitals humanities". [https://fredoc2015.sciencesconf.org/conference/fredoc2015/pages/Acquier\\_FREDOC2015\\_Phonothe\\_que\\_cor](https://fredoc2015.sciencesconf.org/conference/fredoc2015/pages/Acquier_FREDOC2015_Phonothe_que_cor)  
A enlever oui mais j'hésite car il y a une bonne description du travail sur les métadonnées et nous avons peu d'intervention sur le cœur de notre métier mas le reste est sans intérêt

Des compétences à acquérir pour accompagner le partage des données [ne sélectionner que les ressources les plus pertinentes si l'on décide de dire un mot des compétences A décider. Cela rejoint ma question sur une partie consacrée à l'accompagnement.

Les compétences attendues des professionnels de l'IST - Cécile Swiatek (Ligue des bibliothèques européennes de recherche LIBER) - Fredocs 2018 - Démarches innovantes en IST : expérimenter, proposer, (se) réinventer 3-5 octobre 2018, Albi Les métiers de la documentation et de l'IST ont opéré une mue ces cinquante dernières années, au fil des rapides évolutions de la technologie et de nos sociétés. Nous faisons face aujourd'hui à une intéressante équation, dont la combinaison de lignes peut sembler contradictoire. Avec la gestion des données de la recherche et la science ouverte, nous devons effectuer une montée en compétences expertes pour relever un défi technologique – sans tomber dans une dérive techniciste. Avec la logique de la recherche par programmes et par projets, nous formalisons nos politiques et stratégies de documentation et d'IST pour affirmer notre rôle d'appui et de partenaires de la recherche. Ces modèles de pilotage ont l'intérêt de révéler et de porter une vision stratégique de la documentation dans le



milieu de la recherche – mais cela demande de réinventer nos horizons et donc de faire évoluer nos compétences de Leadership. Avec les évolutions des modèles de communication et de publication académiques, couplées aux modalités variables de travail du chercheur actuel – isolé, distant, multi-équipe, fragmenté, électronique, « bibliométré »... – le professionnel de l'IST doit mettre en œuvre de nouvelles formes de médiation, de formation et de communication. Le développement de ces compétences concerne autant sa sphère « métier » que sa qualité de « paraprofessionnel » du chercheur avec lequel il est en relation. Comment articuler ces horizons ? Qu'est-ce qui relève d'un socle de compétences, ou au contraire de l'expertise ? Quelles sont les limites du métier, et comment permettre la reconnaissance de certaines compétences rares, en évolution ou au contraire très générales et parfois laissées de côté ? pdf - <https://fredoc2018.sciencesconf.org/resource/page/id/5> [si nous devons en garder une je propose celle-ci]

#### Évaluation des chercheurs / Qualité de la recherche

La valeur des données de recherche : une nouvelle opportunité pour l'évaluation scientifique ? - Atelier Dialogu'IST 12 juin 2018 Seront mis en exergue le rôle des données scientifiques dans la qualité de la recherche et les pistes qu'elles ouvrent pour l'évaluation de la recherche et des chercheurs. Issus de laboratoires de recherche, de services de documentation et d'organismes de formation, les intervenants proposeront des états des lieux et des témoignages sur ces questions. Basé sur la dynamique d'échanges libres et spontanés, l'atelier en visioconférence permet de donner la parole à tous les acteurs concernés, scientifiques et professionnels de l'information. Au plaisir de vous retrouver nombreux sur l'un de nos site. Le groupe DIALOGU'IST dialoguist chez services.cnrs.fr <http://renatis.cnrs.fr/spip.php?article474> [Pas très explicite sur l'évaluation et pour ce qui est des la qualité des données, d'autres ressources au sein du réseau qualité seront plus pertinentes]

## 2. Concevoir et planifier

Contexte : Définition des tâches, élaboration de plannings, recherche de partenaires, de financement, spécifications, métadonnées, plan de diffusion ...). Les réseaux apportent un appui sur la gestion de projet des méthodologies de conduite de projet: définition des indicateurs, interopérabilité des systèmes, aide à la rédaction des plans de gestion de données, planification du type de données à collecter, identification des nouveaux supports de publications...). Il est nécessaire de prévoir le mode de collecte et de stockage afin d'organiser la traçabilité en amont susceptible de garantir la réutilisation des données

Evaluer les besoins, anticiper les interfaçages nécessaires (avec les utilisateurs ou entre bases de données),

Mettre en place une gestion de projet

Phase de Montage des pprojets : utilisation d'outils de gestion de projets : service de listes de

diffusion partage de documents et de données dans des dossiers partagés en réseau , service de cloud plate forme de gestion de projet : redmine ou autre connaître les possibilités et ressources internes à l'unié et celles fournies par des partenaires extérieurs : Université, CNRS, Renater , etc...

Les aspects de conception et de planification nécessitent de mettre en place une méthodologie de gestion de projet et une analyse des risques Mise en place d'une gestion de projet dans le domaine de la recherche en biologie : [https://qer-2017.sciencesconf.org/data/program/2017\\_ANF\\_tracabilite\\_ferro.pdf](https://qer-2017.sciencesconf.org/data/program/2017_ANF_tracabilite_ferro.pdf)  
THEDRE : Méthode de conduite de la recherche en informatique centrée humain : [https://qer-2017.sciencesconf.org/data/program/2017\\_ANF\\_tracabilite\\_mandran.pdf](https://qer-2017.sciencesconf.org/data/program/2017_ANF_tracabilite_mandran.pdf)  
Outil de Diagnostic EuropeQUA: une méthode pour manager tout type d'activité : [https://qualsimp.sciencesconf.org/data/program/3\\_O](https://qualsimp.sciencesconf.org/data/program/3_O)  
La cartographie des risques pour améliorer les services relatifs à la gestion des contrats et conventions : [https://qualsimp.sciencesconf.org/data/program/17\\_La\\_cartographie\\_des\\_risques\\_pour\\_ame\\_liorer\\_L](https://qualsimp.sciencesconf.org/data/program/17_La_cartographie_des_risques_pour_ameliorer_l)

Disponibilités des infrastructures (lieu, fonctionnalités, capacités) et services

Infrastructure pour l'open science EOSC [https://jcad2019.sciencesconf.org/data/EOSC\\_FranceGrilles\\_Beckman](https://jcad2019.sciencesconf.org/data/EOSC_FranceGrilles_Beckman)  
Infrastructure de grille et de cloud EGI et France Grille Sur EGI : [https://jcad2019.sciencesconf.org/data/EGI\\_upd](https://jcad2019.sciencesconf.org/data/EGI_upd)  
Sur le service de stockage de France Grille : [https://jcad2019.sciencesconf.org/data/Poster\\_FG\\_iRODS\\_2019\\_JC](https://jcad2019.sciencesconf.org/data/Poster_FG_iRODS_2019_JC)  
Sur le cloud France Grille : [https://jcad2019.sciencesconf.org/data/posterJCAD2019\\_Hamrouni\\_vf.pdf](https://jcad2019.sciencesconf.org/data/posterJCAD2019_Hamrouni_vf.pdf)

« Sélectionner, gérer, valoriser et maintenir une plateforme de données en SHS à l'heure du digital turn » : Stéphane Pouyllau (TGIR Huma-Num) - Fredocs2013 - Gestion et valorisation des données de la recherche - 7 au 10 octobre 2013, Aussois [http://renatis.cnrs.fr/IMG/pdf/Pouyllau\\_FREDOC2013.pdf](http://renatis.cnrs.fr/IMG/pdf/Pouyllau_FREDOC2013.pdf)  
[Un peu ancien]

Bases de ressources numériques et services aux chercheurs. Avec ISTEEX et OpenMintedD, l'alliance pour une infrastructure de text-mining - Laurence El Khouri (DIST-CNRS) et Stéphane Schneider (INIST-CNRS) - Fredocs 2018 - Démarches innovantes en IST : expérimenter, proposer, (se) réinventer 3-5 octobre 2018, Albi ISTEEX est une base documentaire qui propose un accès à distance et de manière pérenne à un corpus multidisciplinaire (plus de 21 millions de documents) avec accès systématique au texte intégral. Elle est dotée d'un moteur de recherche puissant, adapté au volume du corpus et aux besoins des scientifiques offrant des facilités d'interrogation et de téléchargement. ISTEEX propose également des services permettant la mise en place de traitements des données : extraction, fouille de textes, production de synthèses documentaires. La présentation enchainera par OpenMinted, qui est une plateforme open source, ouverte et pérenne grâce à laquelle les chercheurs peuvent découvrir, créer, partager et réutiliser des logiciels, des documents et des ressources pour le text-mining, le TALN, l'Extraction d'Information en travaillant à partir de sources documentaires licitement fouillables tels qu'ISTEEX, OPENAIRE et CORE.

pdf - [https://fredoc2018.sciencesconf.org/data/pages/Pres\\_istex\\_visa\\_omtd\\_ElKhouri\\_Schneider.pdf](https://fredoc2018.sciencesconf.org/data/pages/Pres_istex_visa_omtd_ElKhouri_Schneider.pdf)

Positionnement et offre globale de l'INIST dans le contexte IST en évolution - Claire François (INIST) - Fredocs 2018 - Démarches innovantes en IST : expérimenter, proposer, (se) réinventer 3-5 octobre 2018, Albi L'Institut de l'Information Scientifique et Technique, unité de service du CNRS déploie ses activités dans le cadre de son projet d'établissement « Ingénierie des connaissances 2018-2022 » Construit autour de 3 axes cœur « Analyse et fouille de l'information », « Valorisation des données de la recherche », « Accès à l'information scientifique », ce projet s'inscrit dans le contexte de politique nationale de science ouverte. Axe 1 : Analyse et fouille de l'information · Appui au pilotage : aide au recueil des données sur les publications et à la production d'indicateurs bibliométriques correspondants. · Terminologie et fouille de texte : création et de gestion de terminologies scientifiques et développement des approches de fouille de textes Axe 2 : Valoriser les données de la recherche · Outils et services d'accompagnement pour optimiser le partage et l'interopérabilité des données de la recherche (portail OPIDoR) · Formation à distance : plateforme DoRANum intégrant différentes ressources d'autoformation en libre accès sur la gestion et le partage des données de la recherche. Axe 3 : Accès à l'information · BibCnrs : portail multidisciplinaire d'accès aux ressources électroniques · ISTE : plateforme d'accès aux archives scientifiques (plus de 21 millions de documents publiés entre 1400 et 2015) -EzPAARSE/ezMESURE : uiste logicielle de mesure des usages des ressources électroniques pdf - [https://fredoc2018.sciencesconf.org/data/pages/INIST\\_C\\_Francois.pdf](https://fredoc2018.sciencesconf.org/data/pages/INIST_C_Francois.pdf)

Comité pour la Science Ouverte (CoSO) - comment mettre en oeuvre et accompagner la politique de science ouverte - Joanna Janik (DIST-CNRS) - Fredocs 2018 - Démarches innovantes en IST : expérimenter, proposer, (se) réinventer 3-5 octobre 2018, Albi pdf - [https://fredoc2018.sciencesconf.org/data/pages/CoSO\\_J\\_Janik.pdf](https://fredoc2018.sciencesconf.org/data/pages/CoSO_J_Janik.pdf)

Présentation de Doranum - Yvette Lafosse (INIST-CNRS) - Participer à l'organisation du management des données de la recherche : gestion de contenu et documentation des données - 3-6 juil. 2017 Paris (France) ppt - [https://anfdonnees2017.sciencesconf.org/data/pages/Presentation\\_DoRANum\\_ANF\\_0](https://anfdonnees2017.sciencesconf.org/data/pages/Presentation_DoRANum_ANF_0)

« Réseau Quetelet : Banques de données pour les sciences sociales » : Roxane Silberman (TGIR PROGEDO/Réseau Quetelet- Frédocs2013 - Gestion et valorisation des données de la recherche - 7 au 10 octobre 2013, Aussois [http://renatis.cnrs.fr/IMG/pdf/SILBERMAN\\_08102013.pdf](http://renatis.cnrs.fr/IMG/pdf/SILBERMAN_08102013.pdf)

Prévoir les modes de collecte et de stockage

Amorcer un Plan de gestion de données :

Un Data Management Plan (DMP) ou Plan de Gestion de Données (PGD) est un document formalisé explicitant la manière dont seront obtenues, documentées, analysées, disséminées et utilisées les données produites au cours et à l'issue d'un processus ou d'un projet de recherche.

L'initiation du PGD dans cette phase est un préalable à sa mise à jour dans les étapes suivantes. Le PGD doit suivre les évolutions

Plans de gestion de logiciels [https://jcad2019.sciencesconf.org/data/PRESOFT\\_JCAD2019.pdf](https://jcad2019.sciencesconf.org/data/PRESOFT_JCAD2019.pdf)

Présentation et utilité de la plateforme DMP OPIDOR de l'INIST pour rédiger des plans de gestion de données [https://sist18.sciencesconf.org/data/pages/15\\_MC\\_Jacquemot\\_Perbal\\_L\\_Rassinoux\\_OPIDoR.pdf](https://sist18.sciencesconf.org/data/pages/15_MC_Jacquemot_Perbal_L_Rassinoux_OPIDoR.pdf)

Laurent RASSINOX, Institut de l'information scientifique et technique - Marie-Christine JACQUEMOT-PERBAL, Institut de l'information scientifique et technique

Du Plan de Gestion des Données au Datapaper : suivi des données scientifiques tout au long de leur cycle de vie [https://sist18.sciencesconf.org/data/pages/16\\_W\\_Heintz\\_Du\\_plan\\_de\\_gestion\\_des\\_donnees.pdf](https://sist18.sciencesconf.org/data/pages/16_W_Heintz_Du_plan_de_gestion_des_donnees.pdf)  
Wilfried Heintz, UMR 1201 Dynafor

QUIDOZ, Marie-Claude, 2019. Plan de gestion des données. Semaine TEMPO, Sète. 25 septembre 2019. <http://rbdd.cnrs.fr/IMG/pdf/tempo-pgd.pdf?580/01010276848206d6f57d0d6c5d8d93a441f83668>

Créer un plan de gestion des données (DMP) - Marie Puren (INRIA) - Participer à l'organisation du management des données de la recherche : gestion de contenu et documentation des données - 3-6 juil. 2017 Paris (France) ppt - <https://anfdonnees2017.sciencesconf.org/resource/page/id/3>

Accompagner les chercheurs : 5 tutoriels à votre disposition pour la rédaction d'un Plan de Gestion des Données dans le cadre d'Horizon 2020 - Yvette Lafosse, Françoise Cosserat INIST-CNRS - Frédocs 2015 - Mieux accompagner la recherche : réalités d'aujourd'hui et perspectives pour les fonctions IST - 29 septembre - 2 octobre 2015, Saintes-Foy-Lès-Lyon [https://fredoc2015.sciencesconf.org/conference/fredoc2015/pages/Lafosse\\_Atelier\\_FreDoc2015\\_tutosH2020.pdf](https://fredoc2015.sciencesconf.org/conference/fredoc2015/pages/Lafosse_Atelier_FreDoc2015_tutosH2020.pdf)

Structuration générale des données

Caractéristiques des données (formats de fichiers, conventions de nommage ...)

Comment assurer l'interopérabilité ? définir la façon d'harmoniser les données en fonction de leur provenance pour permettre l'interopérabilité.

[[nb: il y a plusieurs phases où parler de l'interopérabilité (concevoir, collecter, traiter etc) elle est à tous les niveaux ... faudra faire attention à ce qu'il n'y ait pas de redites ou redondances]]

Concevoir les bases de données nécessaires

Démarche de gestion des données [... à caser soit dans les infrastructures de recherche au dessus, soit dans l'étape 1 "connaître le paysage de la gestion des données"]

Vers la gestion et le partage des données de la recherche : chercheurs et documentalistes scientifiques des évolutions parallèles - Thierry Beguiristain (LIEC, UMR 7360, CNRS-Université de Lorraine), Marie-Christine Jacquemot-Perbal (INIST-CNRS) - Frédocs 2015 - Mieux accompagner la recherche : réalités d'aujourd'hui et perspectives pour les fonctions IST - 29 septembre - 2 octobre 2015, Saintes-Foy-Lès-Lyon Avec l'essor des technologies et le tournant numérique, chercheurs et documentalistes scientifiques doivent faire évoluer leurs pratiques professionnelles et s'engager vers une science des données ouvertes. Pour relever ce défi qui est non seulement d'ordre

socio-professionnel mais aussi socio-culturel, scientifique, et technique, des chercheurs et des documentalistes choisissent de s'associer et de construire ensemble les nouvelles connaissances et compétences requises. Cette démarche sera présentée à travers le partenariat établi entre l'Observatoire Terre Environnement de Lorraine (OTELo) et l'Inist-CNRS. Celui-ci a été initié dans le cadre d'une étude de cas pour un projet européen SIM4RDM dont l'objectif était d'évaluer les connaissances et compétences des chercheurs en matière de bonnes pratiques de gestion et partage des données de la recherche et, de les y sensibiliser. Actuellement, à l'issue de cette phase de sensibilisation, les documentalistes accompagnent les chercheurs dans l'élaboration et la mise en œuvre de plans de gestion des données. Afin d'assurer ces nouveaux rôles auprès des chercheurs, ils ont construit de nouvelles connaissances en s'appuyant sur leurs connaissances scientifiques et documentaires, sur les ressources en ligne et sur l'expérience d'autres pays plus avancés comme le Royaume Uni ou les Etats-Unis. Ils concrétisent et consolident ces acquis et développent de nouvelles compétences tout en marchant auprès des chercheurs.

[https://fredoc2015.sciencesconf.org/conference/fredoc2015/pages/Beguiristain\\_Jacquemot\\_FreDOC2015.pdf](https://fredoc2015.sciencesconf.org/conference/fredoc2015/pages/Beguiristain_Jacquemot_FreDOC2015.pdf)

« Initiatives régionales : projet Lorrain » : Alain Collignon (INIST/CNRS) - - Frédocs2013 - Gestion et valorisation des données de la recherche - 7 au 10 octobre 2013, Aussois [http://renatis.cnrs.fr/IMG/pdf/1\\_ACO\\_presentation\\_FREDOC\\_final.pdf](http://renatis.cnrs.fr/IMG/pdf/1_ACO_presentation_FREDOC_final.pdf)

« RESIF (données sismologiques et géodésiques) » : Catherine Pequenet (Institut des Sciences de la Terre, Observatoire des Sciences de l'Univers de Grenoble) - Frédocs2013 - Gestion et valorisation des données de la recherche - 7 au 10 octobre 2013, Aussois <http://renatis.cnrs.fr/IMG/pdf/RESIF-FREDOC-SISMO.pdf>

« La gestion des données astronomiques à l'Observatoire Astronomique de Strasbourg » : Soizick Lesteven (Observatoire Astronomique de Strasbourg) - Frédocs2013 - Gestion et valorisation des données de la recherche - 7 au 10 octobre 2013, Aussois <http://renatis.cnrs.fr/IMG/pdf/LESTEVEN.pdf>

### 3. Collecter

Phase de constitution des jeux de données avec leur métadonnées. Les réseaux participent à la constitution des jeux de données, des métadonnées et du choix des référentiels associés, et à la mise en place d'un environnement leur permettant d'assurer le stockage, le «versionning», et l'accessibilité des données et des jeux de données via des bases ou des supports fiables et bien documentés, dans le respect des règles de traitement spécifiques des données personnelles.

nécessité de métadonnées précises pour la description des capteurs de mesures et des dispositifs d'acquisition établissement et documentation des chaînes de collecte plus ou moins complexes : du capteur jusqu'aux disques et aux serveurs où les traitements pourront être établis nécessité de gestion et conduite de projets pour faire travailler ensembles différents acteurs intervenant dans la chaîne de collecte électronique, informaticiens, chercheurs

nécessité de cahier ou tablettes de terrain pour consigner les relevés et métadonnées observées

utilisation de protocoles si possibles normalisés ou standardisés comme SOS “Sensor Observation Service” de l’OGC pour présenter les données de capteurs prévoir, estimer le stockage nécessaire à la collecte de données : travailler en amont avec une équipe informatique e mode projet (gestion de projet) ———

Maîtriser l’acquisition et collecte des données

différents aspects de collectes : par un capteur automatisé, par un personnel de terrain, par un modèle numérique qui produit des données, par une enquête, par des résultats d’analyses, au moyen d’interfaces... élaborer des méthodologies de collecte,

nécessité de métadonnées précises pour la description des capteurs de mesures et des dispositifs d’acquisition établissement et documentation des chaines de collecte plus ou moins complexes : du capteur jusqu’aux disques et aux serveurs où les traitements pourront être établis nécessité de gestion et conduite de projets pour faire travailler ensembles différents acteurs intervenant dans la chaîne de collecte electronicien, informaticiens, chercheurs

nécessité de cahier ou tablettes de terrain pour consigner les relevés et métadonnées observées utilisation de protocoles si possibles normalisés ou standardisés comme SOS “Sensor Observation Service” de l’OGC pour présenter les données de capteurs

---

[...] différents aspects de collectes : par une enquête, par des résultats d’analyses, au moyen d’interfaces... élaborer des méthodologies de collecte, conseiller quant au choix des référentiels de métadonnées (et thésaurus ?) développer les procédures d’intégration des données dans les bases

-Cahiers de laboratoire L’ensemble des données produites par la recherche doit ainsi être répertorié et enregistré dans l’objectif d’une réutilisation potentielle. Nous disposons pour ce faire d’un certain nombre de supports comme les cahiers de laboratoire : [https://qer-2017.sciencesconf.org/data/program/2017\\_ANF\\_tracabilite\\_denis\\_meyere.pdf](https://qer-2017.sciencesconf.org/data/program/2017_ANF_tracabilite_denis_meyere.pdf)

Exemples d’utilisation de cahiers de laboratoire électronique Expérimentation du cahier de laboratoire électronique à l’Inserm : [https://qer-2017.sciencesconf.org/data/program/2017\\_ANF\\_tracabilite\\_dupre.pdf](https://qer-2017.sciencesconf.org/data/program/2017_ANF_tracabilite_dupre.pdf). Cahier de laboratoire électronique BIOVIA ([https://qer-2017.sciencesconf.org/data/program/2017\\_ANF\\_tracabilite\\_biovial.pdf](https://qer-2017.sciencesconf.org/data/program/2017_ANF_tracabilite_biovial.pdf))

-Tablettes de terrain pour consigner les relevés et métadonnées observées

Mettre en oeuvre les bonnes pratiques de traçabilité des activités de recherche : [http://qualite-en-recherche.cnrs.fr/IMG/pdf/guide\\_tracabilite\\_activites\\_recherche\\_gestion\\_connaissances.pdf](http://qualite-en-recherche.cnrs.fr/IMG/pdf/guide_tracabilite_activites_recherche_gestion_connaissances.pdf) AKOUEITE, Ata Franck, 2018. Collecte de données terrain avec un smartphone : Prise en main de Kobotolbox et de Kobocollect. FOSS4G-fr 2018, Marne-la-vallée. 2018. <https://github.com/OSGeo-FOSS4G-fr-2018/blob/master/ateliers/Atelier-kobo.pdf>

BORDÈRES, Serge, 2018. Pi 4x4 : Conception d'une tablette de terrain pour la recherche. In : Atelier "carnets de terrain électroniques" [en ligne]. Montpellier. 2018.[http://rbdd.cnrs.fr/IMG/pdf/borderes\\_atelier2018.pdf](http://rbdd.cnrs.fr/IMG/pdf/borderes_atelier2018.pdf)  
Retour d'expérience sur le montage d'une tablette PI4\*4 (M. Rouan) ([http://rbdd.cnrs.fr/IMG/pdf/pi\\_4\\_4.pdf?578](http://rbdd.cnrs.fr/IMG/pdf/pi_4_4.pdf?578))

Environnements de stockage - Sauvegarder les données

prévoir, estimer le stockage nécessaire à la collecte de données : travailler en amont avec une équipe informatique e mode projet (gestion de projet)

Architectures de stockage traditionnels (ANF 2016) <https://indico.mathrice.fr/event/5/contribution/2/material/1.pdf>

Organiser les données

Choisir les référentiels de métadonnées (thésaurus, ontologies)

Utilisation de standards d'interopérabilité dans la gestion des métadonnées

Pourquoi un catalogue de métadonnées ? Outil de diffusion / valorisation Outil de recherche  
Directive Inspire sur les données géographiques publiques (2007)–Catalogage des données–  
Accès gratuit aux métadonnées

Mise en place de catalogues INSPIRE et de leur alimentation automatique.

[https://sist16.sciencesconf.org/data/pages/14\\_C\\_Bernard\\_J\\_Fabre.pdf](https://sist16.sciencesconf.org/data/pages/14_C_Bernard_J_Fabre.pdf) - Cyril Bernard, CEFE -  
Juliette Fabre, OSU OREME - Olivier Lobry, OSU OREME

INSPIRE un cadre pour mieux partager es données de la recherche : Marc Leobet (Chargé de mission et PCE INSPIRE) - Frédocs2013 - Gestion et valorisation des données de la recherche - 7 au 10 octobre 2013, Aussois Directive inspire - Réutilisation des données du secteur public - Application à la recherche [http://renatis.cnrs.fr/IMG/pdf/Leobet\\_INSPIRE\\_Fredoc2013.pdf](http://renatis.cnrs.fr/IMG/pdf/Leobet_INSPIRE_Fredoc2013.pdf)

Format et métadonnées

Format et métadonnées - Catherine Morel-Pair - Participer à l'organisation du management des données de la recherche : gestion de contenu et documentation des données - 6-8 juil. 2016 Paris (France) Vidéo - <https://youtu.be/obGDFrXyBiU?t=3> ppt - <https://anfdonnees2016.sciencesconf.org/resource/pag>

Format et métadonnées - Catherine Morel-Pair - Participer à l'organisation du management des données de la recherche : gestion de contenu et documentation des données - 6-8 juil. 2016 Paris (France) Vidéo - <https://youtu.be/obGDFrXyBiU?t=3> ppt - <https://anfdonnees2016.sciencesconf.org/resource/pag>

Les métadonnées dans un DMP- Marie Puren (INRIA) - Participer à l'organisation du management des données de la recherche : gestion de contenu et documentation des données - 3-6 juil. 2017 Paris (France) [https://anfdonnees2017.sciencesconf.org/data/pages/20170706\\_dmp\\_metadonnees\\_puren\\_1.pdf](https://anfdonnees2017.sciencesconf.org/data/pages/20170706_dmp_metadonnees_puren_1.pdf)

Utiliser des Standards d'interopérabilité -

- Protocoles interopérables (en sciences de l'environnement)

les gestionnaires de données environnementales mettent en place des chaînes de collecte de données provenant de capteurs de terrains, ou de modèles numériques. Ils se préoccupent de l'utilisation de normes interopérables dans les protocoles d'échange et dans les formats de données

Cette partie traite de l'Utilisation des standards d'interopérabilité en sciences de l'environnement  
A cet effet l'OGC (open géospatial consortium) publie différents standards d'interopérabilité, dont SWE "Sensor Web Enablement", qui permet de présenter des données de capteur de manière standardisée et interopérable [à détailler]

Ces protocoles standards de l'OGC ont été présentés lors du Séminaire du réseau SIST2015 à l'OSU Pytheas de Marseille

Présentation des différents protocoles interopérables proposés par l'OGC <https://sist15.sciencesconf.org/data/p>  
François André OMP Toulouse

Présentation du standard "SWE" <https://sist15.sciencesconf.org/program> (mettre le lien exact vers le fichier) christoph Stasch 52North

Le programme Seadatanet vise à élaborer et mettre en place un portail Européen d'accès aux données marines. Il se base sur de nombreux standards rendants les données FAIR, cherchable, Interopérables et Réutilisables Seadatanet est un exemple d'envergure pour la mise en place de standards d'interopérabilité à l'échelle Européenne

Présentation du projet SeaDataNet, interopérabilité à l'échelle pan-européenne <https://sist15.sciencesconf.org/>  
(mettre le lien exact vers le fichier) Michèle Fichaut, Systèmes d'Informations Scientifiques pour la MER

Portail Web d'accès aux données de l'observatoire AMMA-CATCH et mise en oeuvre des standards d'échange des données OGC  
<https://sist15.sciencesconf.org/program> (mettre le lien exact vers le fichier) Véronique CHAFFARD, Institut de Recherche pour le Développement, Laboratoire d'étude des transferts en hydrologie et environnement

SI-TEC-PSO: retour d'expérience sur le système d'information dédié capteurs et reconstitution de séries temporelles de ReefTEMPS, le réseau de suivi de température des eaux côtières dans la région du Pacifique Sud et Sud-Ouest <https://sist15.sciencesconf.org/program> (mettre le lien exact vers le fichier) Sylvie Fiat, et Régis Hocdé, Institut de Recherche pour le Développement

De la définition au déploiement de standards d'interopérabilité : retour d'expérience de la Direction des Systèmes d'Information (DSI) du BRGM <https://sist15.sciencesconf.org/program> (mettre le lien exact vers le fichier) Grellet Sylvain, BRGM - Stéphane Loigerot, BRGM



Le projet Dat@OSU de gestion et valorisation des données de la recherche

[https://sist16.sciencesconf.org/data/pages/13\\_B\\_Debray.pdf](https://sist16.sciencesconf.org/data/pages/13_B_Debray.pdf)

- Bernard Debray, Univers, Transport, Interfaces, Nanostructures, Atmosphère et environnement, Molécules

Le protocole SOS (Sensor observation service) de l'OGC permet de présenter de manière standardisée les données issues de capteurs de terrain. Certains logiciels comme 52North et istSOS permettent de gérer les données de capteurs dans une BD internet et de fournir les données de capteurs de manière standardisée via le protocole SOS. Ces protocoles standards de l'OGC ont été présentés lors du Séminaire du réseau SIST2015 à l'OSU Pytheas de Marseille <https://sist15.sciencesconf.org/program>

Traiter les données environnementales revient fréquemment à savoir les stocker et les gérer dans des BD relationnelles puis à les représenter sous forme de graphes 1D ou de cartes 2D - Lorsque les valeurs sont acquises en 1 point en continue, on représente leur évolution en fonction du temps, on parle alors de "séries temporelles" (timeseries) - Lorsque les valeurs sont acquises en 1 point le long d'un axe vertical profondeur ou altitude, on représente leur évolution en fonction du temps, on parle alors de "profils verticaux" (profiles)

Présentation du logiciel istSOS <https://sist15.sciencesconf.org/program> (mettre le lien exact vers le fichier) Massimiliano Canata

Collecte de mesures météorologiques à l'aide d'un système autonome : exemple de la métropole rennaise (Zone Atelier Armorique)

<https://sist15.sciencesconf.org/program> (mettre le lien exact vers le fichier) Alban Thomas, Littoral, Environnement, Télédétection, Géomatique

Distribution et visualisation de données avec Thredds, exemples d'utilisation au SEDOO <https://sist15.sciencesconf.org/program> (mettre le lien exact vers le fichier) Guillaume Brissebrat, Service de données de l'OMP

Base de données «Observation» à la Station Biologique de Roscoff : retours d'expérience. <https://sist15.sciencesconf.org/program> (mettre le lien exact vers le fichier) Mark Hoebeke, Station biologique de Roscoff [Roscoff]

Outil web interactif de visualisation et validation de séries temporelles. <https://sist15.sciencesconf.org/program> (mettre le lien exact vers le fichier) Olivier LOBRY, Observatoire de REcherche Méditerranéen de l'Environnement

développer les procédures d'intégration des données dans les bases

[RBdD] (à répartir) L'ANF « Interfacer les outils mobiles avec son système d'information » portait principalement sur ces questions : Outils nomades : définir ses besoins. Tour d'horizon des applis embarquées et retour d'expérience (C. Plumejeaud, S.Ladet) ([http://rbdd.cnrs.fr/IMG/pdf/anf\\_rbdd\\_2019\\_outil](http://rbdd.cnrs.fr/IMG/pdf/anf_rbdd_2019_outil))

besoins\_2juin2019.pdf?577/a94b104f974483c5dbda6c4939ef0db9422bd1dd) Présentation de la solution ODK (MC. Quido) ([http://rbdd.cnrs.fr/IMG/pdf/deployer\\_avec\\_odk.pdf?569/9c7a0e508a744bc2ddd7](http://rbdd.cnrs.fr/IMG/pdf/deployer_avec_odk.pdf?569/9c7a0e508a744bc2ddd7)) Comment faire un formulaire avec XLSForm ? (PY. Arnould) ([http://rbdd.cnrs.fr/IMG/pdf/xlsform\\_anf2019-2.pdf?575/d45519bbb0360384f14bcf6f4c072313eb7c60a5](http://rbdd.cnrs.fr/IMG/pdf/xlsform_anf2019-2.pdf?575/d45519bbb0360384f14bcf6f4c072313eb7c60a5)) Mobile Atlas Creator (A. Cheylan) (<http://rbdd.cnrs.fr/IMG/pdf/mobac.pdf?572/f49254091f49630e2b2088ec36964cbd931e0278>) Intégrer les données dans sa base métier (MC. Quido) ([http://rbdd.cnrs.fr/IMG/pdf/integrer\\_donnees.pdf?570/20](http://rbdd.cnrs.fr/IMG/pdf/integrer_donnees.pdf?570/20)) UUID avec PostgreSQL : Pourquoi ? Comment ? (N. Raidelet) ([http://rbdd.cnrs.fr/IMG/pdf/uuid\\_postgres.pdf?40](http://rbdd.cnrs.fr/IMG/pdf/uuid_postgres.pdf?40)) Les + et les - de la solution ODK (MC. Quido & PY. Arnould & les stagiaires) ([http://rbdd.cnrs.fr/IMG/pdf/bilan\\_odk](http://rbdd.cnrs.fr/IMG/pdf/bilan_odk)) Référence sur ODK (MC. Quido) (<http://rbdd.cnrs.fr/IMG/pdf/bibliographie.pdf?571/bbecc7bd883e4751efb6ccb>) Outils nomades : validation des données (C. Plumejeaud) ([http://rbdd.cnrs.fr/IMG/pdf/anf\\_rbbd\\_2019\\_outils\\_n](http://rbdd.cnrs.fr/IMG/pdf/anf_rbbd_2019_outils_n)) Retour terrain : la délicate question de l'intégration des données (PY. Arnould) (<http://rbdd.cnrs.fr/IMG/pdf/anf2>)

#### 4. Traiter

Contexte: Cette étape du cycle de vie des données est destinée à Filtrer, regrouper, choisir les données pertinentes, reformater, gérer les métadonnées : Les réseaux participant au traitement des jeux de données via la mise en place de procédures ou logiciels assurant la qualité des données et la qualité des traitements (modélisation et reformatage des données ; mise en base de données ; mise à disposition de moyens informatiques; aide aux choix de la solution la plus adaptée, détection des erreurs de saisie ou des incohérences relatives aux modèles de données, mise en place d'un contrôle qualité)

Chaines et méthodes de traitement de données

Logiciels et plateforme logicielles de traitement

Introduction au langage de programmation Julia et son utilisation dans le cadre du traitement de données. Journée Julia, Map Reduce, janvier 2019 <https://calcul.math.cnrs.fr/2019-01-journee-julia.html#collapseSupports1> Xavier Vasseur, ISAE

Python et l'écosystème disponible pour la data science <https://calcul.math.cnrs.fr/2017-12-journee-python-data.html>

Traitement de fichiers - Format de fichiers interopérables

mise dans des formats interopérables NetCDF, ODV, HDF, etc..

Copier les succès et rester simple (AMEO) : Mise à disposition de sorties de modèles climatiques avec un NAS, THREDDS et ERDDAP.

[https://sist16.sciencesconf.org/data/pages/11\\_T\\_Valero\\_F\\_Bongat.pdf](https://sist16.sciencesconf.org/data/pages/11_T_Valero_F_Bongat.pdf) - Thierry VALERO, Institut de Recherche pour le Développement, Laboratoire d'Océanographie et du Climat : Expérimentations et Approches Numérique

[https://sist16.sciencesconf.org/data/pages/12\\_R\\_Hocde.pdf](https://sist16.sciencesconf.org/data/pages/12_R_Hocde.pdf) Réseau d'observation du Pacifique

Sud 'ReefTEMPS' : évolutions fonctionnelles et optimisation d'un système d'information dédié capteurs et reconstitution de séries temporelles

[https://sist16.sciencesconf.org/data/pages/12\\_R\\_Hocde.pdf](https://sist16.sciencesconf.org/data/pages/12_R_Hocde.pdf)

- Régis Hocdé, Institut de Recherche pour le Développement

Métrologie des équipements

La confiance dans la qualité d'une recherche consiste à établir et vérifier que les différentes étapes d'une étude peuvent être répétées en obtenant le même résultat par des chercheurs différents à des moments différents. De ce fait, il convient d'avoir une totale maîtrise des équipements de traitement : Métrologie des équipements : [https://qualsimp.sciencesconf.org/data/program/9\\_Trac\\_abilite\\_de](https://qualsimp.sciencesconf.org/data/program/9_Trac_abilite_de)

Contrôle Qualité des données - Mise en place de procédures qualité

Plusieurs présentations et ateliers sur ce thème ont eu lieu lors de l'ANF « Sciences des données : un nouveau challenge pour les métiers liés aux bases de données » (<http://rbdd.cnrs.fr/spip.php?article288>)

- du 5 au 7 novembre 2018 à Sète. En particulier l'Atelier qualité des données dont les travaux portaient sur les questions suivantes : Quelles sont les différentes notions de qualité des données ? Comment contrôler la qualité des données dans la BDD : avant ou pendant l'insertion de données Faut-il automatiser le contrôle de la qualité dans les bases ? Quels sont les outils disponibles et comment les utiliser ? L'Introduction méthodologique et terminologique ([http://rbdd.cnrs.fr/IMG/pdf/qualite\\_des\\_donnees\\_plumejeaud\\_2018\\_04112018.pdf?517/365a13edab604bd07](http://rbdd.cnrs.fr/IMG/pdf/qualite_des_donnees_plumejeaud_2018_04112018.pdf?517/365a13edab604bd07)) a été suivie d'un cours et de TP portant sur "OpenRefine pour traiter son fichier d'entrée" (<http://rbdd.cnrs.fr/IMG/pdf/openrefinecours.pdf?518/a69ce451abd02003a0e96957e39828e0f2e9f2ee>, <http://rbdd.cnrs.fr/IMG/pdf/openrefinedoc.pdf?519/a6de5321fdbedeec29da6cc8b82250d02937ddeb>, <http://rbdd.cnrs.fr/IMG/zip/exos.zip?520/e51f82826431b71f767e4347fd57716fa9175664>) On pourra aussi de référer à quelques présentations de l'ANF « Interfacer les outils mobiles avec son système d'information » citée au chapitre 3 (Collecter) : Outils nomades : validation des données (C. Plumejeaud) ([http://rbdd.cnrs.fr/IMG/pdf/anf\\_rbdd\\_2019\\_outils\\_mobiles\\_tp\\_qualite.pdf?573/e1425561fd10](http://rbdd.cnrs.fr/IMG/pdf/anf_rbdd_2019_outils_mobiles_tp_qualite.pdf?573/e1425561fd10)) Retour terrain : la délicate question de l'intégration des données (PY. Arnould) (<http://rbdd.cnrs.fr/IMG/pdf/anf2>)

BLONDEL, Emmanuel, 2018. Ecrire et Lire des métadonnées avec la librairie R geometa. In : Atelier « Métadonnées et R ». Montpellier. 2018. [http://rbdd.cnrs.fr/IMG/pdf/workshop\\_r\\_metadata\\_agropolis\\_-\\_geometa.pdf?504/f5cc31589976b1cd1fc18d406a547ee18122c0e7](http://rbdd.cnrs.fr/IMG/pdf/workshop_r_metadata_agropolis_-_geometa.pdf?504/f5cc31589976b1cd1fc18d406a547ee18122c0e7)

[resinfo] Traitement de grosses masses de données (BigData), regroupement de données, filtrage et extraction des données pertinentes (ANF 2016) <https://indico.mathrice.fr/event/5/contribution/19/material/s>

-Curation des données [il faudra expliciter le terme et les concepts svp]

« La gestion des données astronomiques à l'Observatoire Astronomique de Strasbourg » : Soizick Lesteven (Observatoire Astronomique de Strasbourg) - Frédocs2013 - Gestion et valorisation des données de la recherche - 7 au 10 octobre 2013, Aussois <http://renatis.cnrs.fr/IMG/pdf/LESTEVEN.pdf>

« De quelques défis spécifiques de la curation numérique des données en SHS : petite incursion dans l'univers de l'édition critique de sources au format TEI » : Emmanuelle Morlock (HiSoMa) - - Frédocs2013 - Gestion et valorisation des données de la recherche - 7 au 10 octobre 2013, Aussois [A conserver je pense] <http://renatis.cnrs.fr/IMG/pdf/emma-morlock-fredocs-2013.pdf>

## 5. Analyser

Contexte: La participation à l'analyse des données se réalise au travers d'un accompagnement sur la mise en place de procédures qualité, par l'optimisation des calculs ou des temps d'accès, ainsi que dans le choix et la mise en œuvre des techniques et outils d'analyse et éventuellement dans leur conception, en fonction des besoins (formats, interopérabilité, performance, ergonomie, visualisation, travail collaboratif...)

Traitement de grosses masses de données (BigData), regroupement de données, filtrage et extraction des données pertinentes (ANF 2016) <https://indico.mathrice.fr/event/5/contribution/19/material/slides/0.p>

Nouveaux outils d'analyse des masses de données (IA) (JoSy 2018) <https://indico.mathrice.fr/event/130/contrib>

Chaîne et méthode d'analyse des données - Utilisation d'outils d'analyse

- Optimisation des calculs ou des temps d'accès

[RENATIS] « Le défi MASTODONS : un instrument pour la gestion et l'exploitation de grandes masses de données » : Mokrane Bouzeghoub (DAS INS2I, Mission pour l'interdisciplinarité, CNRS, projet Mastodons) - Frédocs2013 - Gestion et valorisation des données de la recherche - 7 au 10 octobre 2013, Aussois <http://renatis.cnrs.fr/IMG/pdf/Mastodons-Fredoc-Aussois.pdf>

- Traitement sémantiques/ linguistiques

« Des technologies sémantiques pour l'information scientifique et technique » : Claire Nedellec et Agnès Girard (INRA) - Frédocs2013 - Gestion et valorisation des données de la recherche - 7 au 10 octobre 2013, Aussois [http://renatis.cnrs.fr/IMG/pdf/Nedellec\\_Girard\\_Fredoc-v4.pdf](http://renatis.cnrs.fr/IMG/pdf/Nedellec_Girard_Fredoc-v4.pdf)

« Annotation des Bulletins de santé du végétal (BSV) et interrogation » : Fabien Amarger (IRIT-IRSTEA) - - Frédocs2013 - Gestion et valorisation des données de la recherche - 7 au 10 octobre 2013, Aussois [http://renatis.cnrs.fr/IMG/pdf/FreDoc\\_Irstea.pdf](http://renatis.cnrs.fr/IMG/pdf/FreDoc_Irstea.pdf)

Nouveaux lieux, nouvelles compétences, nouveaux usages. Quels critères de qualités dans de nouveaux espaces de circulation du savoir ? Atelier Dialogu'IST 5 novembre 2017 L'objectif de cet atelier est de montrer au cours de celui-ci qu'au travers de retours d'expériences, d'échanges, les professionnels IST collaborent avec des scientifiques, informaticiens, hors des bibliothèques.

Nous avons choisi de mettre en avant cette mission spécifique concernant la démarche qualité dans le cadre du traitement des données scientifiques, ou encore dans l'analyse et l'extraction des données. Nous voulons faire ressortir au cours de cette journée que l'acquisition de ces compétences par des formations, ou avec l'aide des réseaux métiers,... pourra permettre aux

professionnels IST de répondre aux besoins des scientifiques ou de sensibiliser les chercheurs au respect de certains critères <http://renatis.cnrs.fr/spip.php?article359>

Un partage d'expérience sur la réalisation d'une application d'exploration de données [donner des précisions] [sist] Filtrage interactif de données multi-dimensionnelles <https://sist16.sciencesconf.org/data/pag>  
Patrick Brockmann, Laboratoire des Sciences du Climat et de l'Environnement [Gif-sur-Yvette]

Site Web de diffusion des données "Sahelian Dust Transect"  
[https://sist16.sciencesconf.org/data/pages/10\\_A\\_Campos.pdf](https://sist16.sciencesconf.org/data/pages/10_A_Campos.pdf) - André CAMPOS, Laboratoire inter-universitaire des systèmes atmosphériques

Gestion des données de cytométrie en flux  
[https://sist16.sciencesconf.org/data/pages/06\\_S\\_Lahbib.pdf](https://sist16.sciencesconf.org/data/pages/06_S_Lahbib.pdf) - SOUMAYA LAHBIB, MIO - Mathilde DUGENNE, - Melilotus Thyssen, Institut méditerranéen d'océanologie, - Maurice Libes OSU Pytheas

L'usage d'outils, de données et de webservices interopérables pour la cartographie des formations à silex en France : l'expérience du PCR « Réseau de lithothèques " <https://sist16.sciencesconf.org/data/pag>  
- Christophe Tufféry, Cités, Territoires, Environnement et Sociétés, Institut National de Recherches Archéologiques Préventives

Réseau d'observation du Pacifique Sud 'ReefTEMPS' : évolutions fonctionnelles et optimisation d'un système d'information dédié capteurs et reconstitution de séries temporelles  
[https://sist16.sciencesconf.org/data/pages/12\\_R\\_Hocde.pdf](https://sist16.sciencesconf.org/data/pages/12_R_Hocde.pdf)  
- Régis Hocdé, Institut de Recherche pour le Développement

Contrôle Qualité des données - Mise en place de procédures qualité

[QeR] La confiance dans la qualité d'une recherche consiste à établir et vérifier que les différentes étapes d'une étude peuvent être répétées en obtenant le même résultat par des chercheurs différents à des moments différents. De ce fait, il convient d'avoir une totale maîtrise des équipements d'analyse : [https://qualsimp.sciencesconf.org/data/program/9\\_Trac\\_abilite\\_des\\_donnees\\_de\\_l](https://qualsimp.sciencesconf.org/data/program/9_Trac_abilite_des_donnees_de_l)

[sist] Suivi de la qualité des mesures de réseaux d'observations océanographiques <https://sist16.sciencesconf.org>  
- Philippe Téchiné, Laboratoire d'études en Géophysique et océanographie spatiales

## 6. Archiver, pérenniser (préserver et archiver)

Contexte: L'essentiel de l'accompagnement des réseaux sur cette étape se formalise par l'intégration des notions de pérennisation des données, de traçabilité et d'intégrité des données ainsi que des outils permettant de les exploiter. Les réseaux sont également amenés à participer à la définition des politiques de sauvegarde et d'archivage, et à la mise en place des procédures associées dans les collectifs de recherche.

Sélectionner les données relevantes à pérenniser [quelles sont les données qui nécessitent d'être

pérennisées sur le long terme ?]

Reproductibilité des données

<https://calcul.math.cnrs.fr/2019-01-journee-julia.html> Campagnes de calcul reproductibles

<https://calcul.math.cnrs.fr/2016-05-mini-symposium-canum.html> recherche reproductible

Enjeux de la reproductibilité (JoSy 2019) <https://indico.mathrice.fr/event/165/contribution/1/material/slides/0>.

<https://webcast.in2p3.fr/video/les-enjeux-et-defis-de-la-recherche-reproductible>

Traçabilité de la donnée [RBdD] Plusieurs présentations et ateliers sur ce thème ont eu lieu lors de l'ANF « Sciences des données : un nouveau challenge pour les métiers liés aux bases de données » (<http://rbdd.cnrs.fr/spip.php?article288>) - du 5 au 7 novembre 2018 à Sète. En particulier lors de l'Atelier traçabilité des données Présentation de l'atelier traçabilité des données [http://rbdd.cnrs.fr/IMG/pdf/atelier\\_tracabilite.pdf?523/29abaadfb5e2e0fff8aed53afd88d7aad1ded34f](http://rbdd.cnrs.fr/IMG/pdf/atelier_tracabilite.pdf?523/29abaadfb5e2e0fff8aed53afd88d7aad1ded34f)

e maj : un cas d'utilisation pour tracer les données [http://rbdd.cnrs.fr/IMG/pdf/emaj.2.3.1\\_overview\\_fr.pdf?521/](http://rbdd.cnrs.fr/IMG/pdf/emaj.2.3.1_overview_fr.pdf?521/) et [http://rbdd.cnrs.fr/IMG/pdf/tp\\_e-maj.pdf?522/cbfcf7b13ae9a4d8d20ec495c1ef5ea1d09e0a3f](http://rbdd.cnrs.fr/IMG/pdf/tp_e-maj.pdf?522/cbfcf7b13ae9a4d8d20ec495c1ef5ea1d09e0a3f)

HEINTZ, Wilfried, 2018. Gestion pérenne des données scientifiques : du plan de gestion de données au datapaper. In : Storage Day 2018 [en ligne]. Paris. 16 janvier 2018.

[http://rbdd.cnrs.fr/IMG/pdf/sd2018\\_datamanagement\\_wheintz.pdf?435/8949f16992fdffdf7b7ef53258e71f95](http://rbdd.cnrs.fr/IMG/pdf/sd2018_datamanagement_wheintz.pdf?435/8949f16992fdffdf7b7ef53258e71f95)

LE BA, Nathalie, 2018. Questions juridiques autour de l'ouverture des données. In : ANF « Sciences des données : un nouveau challenge pour les métiers liés aux bases de données » [en ligne]. Sète.

11 mai 2018. [http://rbdd.cnrs.fr/IMG/pdf/20181107\\_questions\\_juridiques\\_autour\\_de\\_l\\_ouverture\\_des\\_donnee](http://rbdd.cnrs.fr/IMG/pdf/20181107_questions_juridiques_autour_de_l_ouverture_des_donnee)

Pérenniser les données - Archivage pérenne

Pérenniser les expériences (JoSy 2019) <https://indico.mathrice.fr/event/165/contribution/3/material/slides/0.p>

<https://webcast.in2p3.fr/video/la-reproductibilite-au-service-de-la-biologie-computationnelle>

nécessité d'Archivage pérenne, politique du CINES Patrimoine scientifique en danger : des solutions d'avenir existent déjà.

[https://sist16.sciencesconf.org/data/pages/18\\_M\\_Massol.pdf](https://sist16.sciencesconf.org/data/pages/18_M_Massol.pdf) - Marion MASSOL, CINES

Software Heritage, préservation des codes sources [https://jcad2019.sciencesconf.org/data/2019\\_10\\_10\\_jcadSo](https://jcad2019.sciencesconf.org/data/2019_10_10_jcadSo)

« La préservation des données scientifiques » : Cristinel Diaconu (Centre de Physique des Particules de Marseille CNRS/IN2P3 et Aix-Marseille Université)- Frédocs2013 - Gestion et valorisation des données de la recherche - 7 au 10 octobre 2013, Aussois [http://renatis.cnrs.fr/IMG/pdf/DIACONU\\_DPHEP\\_PR](http://renatis.cnrs.fr/IMG/pdf/DIACONU_DPHEP_PR)

De l'écoute à l'archivage à long terme, méthodologie du traitement d'archives sonores (pdf) - Françoise Acquier (CRESSON, Ecole d'architecture Grenoble), Ariane Néroutidis, Marine Soubrié (Phonothèque Aix en Provence) - Participer à l'organisation du management des données de la recherche : gestion de contenu et documentation des données - 3-6 juil. 2017 Paris (France) ppt - <https://anfdonnees2017.sciencesconf.org/resource/page/id/3>

## 7. Publier et diffuser

Contexte: Etape finale d'un projet de recherche impliquant une gestion des données, l'accompagnement des réseaux s'exerce sur la documentation des données et leurs formats d'exploitation, pour en permettre l'accès et en assurer la reproductibilité ou la répliquabilité. Ainsi les réseaux travaillent sur l'ensemble des informations (données, modes opératoires, échantillons, publications, visualisation et interfaces graphiques) nécessaires à la mise en œuvre des supports de diffusion et de valorisation les plus pertinents en rapport avec l'objectif du projet initial

Documenter les actions - finaliser le PGD

\*Vérifier corriger Publier les métadonnées

utilisation de standard d'interopérabilité dans la publication des métadonnées

Communiquer - Sensibiliser les acteurs

Publier les données - initiatives nationales - Textes réglementaires Plan national science ouverte <https://www.ouvrirlascience.fr/plan-national-pour-la-science-ouverte/> ouvrir la science <https://www.ouvrirlascience.fr/college-donnees-de-la-recherche-actions-2020/>

RDA <https://www.rd-alliance.org/data-science> directive Européenne INSPIRE

Les pôles de données en Environnement La mission de l'IR Data Terra consiste à organiser de manière intégrée l'accès aux données, en mettant à disposition les données, les produits et des services relatifs à l'observation du système Terre, via les pôles de données et de services existants : <https://www.data-terra.org/> La mission principale des 4 Pôles (ODATIS, AERIS, ForM@Ter & Theia) de l'IR Data Terra est de mettre à disposition des données, des produits, des logiciels, des outils et/ou des services destinés en premier lieu à la communauté scientifique française dans le cadre de ses recherches sur le système Terre. Les informations proposées par les pôles de données sont aussi fondamentales pour la mise en œuvre des politiques publiques. En permettant de mieux comprendre la structure et le fonctionnement du système Terre, les travaux utilisant ces données ont un impact socio-économique important dans des domaines tels que les risques naturels, le changement climatique, les ressources minérales ou les ressources en eau. Dans ce contexte, les pôles servent aussi la communauté internationale (missions satellites, réseaux d'observations internationaux, partenariats pour le développement). Données océanographiques <https://www.odatis-ocean.fr/> Données atmosphériques : <https://www.aeris-data.fr/> Données terre solide: <https://www.poleterresolide.fr/> Données surfaces continentales : <https://www.theia-land.fr/>

[RBdD] (voir ci on laisse ça ici ou dans le paragraphe juridique) Plusieurs présentations et ateliers sur ce thème ont eu lieu lors de l'ANF « Sciences des données : un nouveau challenge pour les métiers liés aux bases de données » (<http://rbdd.cnrs.fr/spip.php?article288>) - du

5 au 7 novembre 2018 à Sète. En particulier dans les présentations suivantes : Questions juridiques autour de l'ouverture des données, Intervenante : Natalie LE BA (DAJ du CNRS, [http://rbdd.cnrs.fr/IMG/pdf/20181107\\_questions\\_juridiques\\_autour\\_de\\_l\\_ouverture\\_des\\_donnees\\_rbdd.pdf?](http://rbdd.cnrs.fr/IMG/pdf/20181107_questions_juridiques_autour_de_l_ouverture_des_donnees_rbdd.pdf?)

[Resinfo]

Virtual Imaging Platform : pour une science ouverte et reproductible (JoSy 2019) <https://indico.mathrice.fr/event/>  
<https://webcast.in2p3.fr/video/la-plateforme-web-virtual-imaging-platform-pour-une-science-ouverte-et-reproductible>

Publier avec des Identifiants pérennes (DOI)

Les données doivent disposer d'un identifiant perenne pour qu'elles soient trouvables, visibles et accessibles au même titre que les publications.

Qu'elles soient plus facilement citables et que l'on puisse les lier aux publications ou à tout autre produit de recherche. Il concourt à l'interopérabilité des données. Attention la pérennité est purement une question de service et n'est pas inhérente à un objet, ni conféré par une syntaxe de nommage particulier. La pérennité du lien vers la localisation de la ressource est de la responsabilité du déposant ou du créateur de l'identifiant.

Pour en savoir plus sur les identifiants pérennes : <https://dorum.fr/identifiants-perennes-pid/>;  
et le DOI de DataCite : <https://seminaire.inrae.fr/data/content/download/3449/36374/version/1/file/03+Yahia+c>  
nécessité de publier en identifiant les jeux de données par des "DOI": DOI de DataCite : un système d'identification pour valoriser vos données de la recherche présentation : [https://sist16.sciencesconf.org/data/pages/15\\_M\\_Yahia.pdf](https://sist16.sciencesconf.org/data/pages/15_M_Yahia.pdf) video : <https://nuage.osupytheas.fr/s/hxjuXQwST6>  
DOI-Yahia-Inist-SIST2016.mp4

Mohamed Salah Yahia, Institut de l'information scientifique et technique du CNRS

Mise en place d'un DOI sur les données d'un réseau d'observations océanographiques  
[https://sist16.sciencesconf.org/data/pages/16\\_P\\_Techine.pdf](https://sist16.sciencesconf.org/data/pages/16_P_Techine.pdf)

- Philippe Téchiné, Laboratoire d'études en Géophysique et océanographie spatiales

Création de DOI sur les données et produits grillés du Service National d'Observation SSS  
[https://sist18.sciencesconf.org/data/pages/14\\_P\\_Techine\\_DOI\\_sur\\_les\\_donnees\\_du\\_SNO\\_SSS.pdf](https://sist18.sciencesconf.org/data/pages/14_P_Techine_DOI_sur_les_donnees_du_SNO_SSS.pdf)  
Philippe Téchiné, Laboratoire d'études en Géophysique et océanographie spatiales

Retour d'expérience sur l'attribution de DOI à l'OSU OREME

[https://sist16.sciencesconf.org/data/pages/17\\_O\\_Lobry.pdf](https://sist16.sciencesconf.org/data/pages/17_O_Lobry.pdf)

Juliette Fabre, OSU OREME - Olivier Lobry, OSU OREME - Séminaire SIST2016 à Montpellier

Création de DOI sur les données et produits grillés du Service National d'observation SSS  
[http://rbdd.cnrs.fr/IMG/pdf/techine\\_doi\\_sss.pdf?525/cde234e7d87d1df64fffd20e8765ab33ca37f7b9](http://rbdd.cnrs.fr/IMG/pdf/techine_doi_sss.pdf?525/cde234e7d87d1df64fffd20e8765ab33ca37f7b9)

Utilisation d'un outil de catalogage normalisé ISO19139 comme GeoNetwork pour constituer une



“landing page” pour un D.O.I [https://sist19.sciencesconf.org/data/pages/SIST19\\_A\\_BATTAIS.pdf](https://sist19.sciencesconf.org/data/pages/SIST19_A_BATTAIS.pdf)

Patrimoine scientifique en danger : des solutions d'avenir existent déjà.

[https://sist16.sciencesconf.org/data/pages/18\\_M\\_Massol.pdf](https://sist16.sciencesconf.org/data/pages/18_M_Massol.pdf) - Marion MASSOL, CINES

"DataCite : identifiants pérennes pour le partage des données » : Herbert Gruttemeier (INIST/CNRS) - Frédocs2013 - Gestion et valorisation des données de la recherche - 7 au 10 octobre 2013, Aussois [http://renatis.cnrs.fr/IMG/pdf/DataCite\\_FreDoc.pdf](http://renatis.cnrs.fr/IMG/pdf/DataCite_FreDoc.pdf)

Déposer/Publier dans des entrepôts institutionnels

[RBDD] Les entrepôts de données, Intervenant : Laurent PELLETIER, <http://rbdd.cnrs.fr/IMG/pdf/20181029-jrbdd-presentationentrepotdonnees.pdf?516/9dc0c4b9f755d121305cf83c440d1d8faed96d73>

Les entrepôts de données : pierre angulaire du partage des données de la recherche - Ester Dzale Yeumo (INRA) - Participer à l'organisation du management des données de la recherche : gestion de contenu et documentation des données - 6-8 juil. 2016 Paris (France)

ppt - <https://anfdonnees2016.sciencesconf.org/resource/page/id/1>

Entrepôts de données - Sylvie Cocard (INRA) - Participer à l'organisation du management des données de la recherche : gestion de contenu et documentation des données - 3-6 juil. 2017 Paris (France) Dans un contexte de sciences ouvertes, les acteurs de la recherche s'accordent aujourd'hui pour considérer les données de la recherche comme des produits de la recherche et appellent à mieux les gérer et à les partager. En France, la loi CADA modifiée par la loi Valter le 28 décembre 2012 incite à mettre les données de la recherche à disposition sous format ouvert et librement réutilisable. Les enjeux liés à la gestion et au partage des données de la recherche nécessitent des outils appropriés. Mais qu'est-ce qu'un entrepôt de données et quelles en sont les principales caractéristiques ? Comment les entrepôts de données contribuent-ils à la gestion et au partage des données ? Qu'est-ce qui différencie un entrepôt de données d'une base de données classique dans le contexte de l'ouverture des données ? Quels services peut-on attendre d'un entrepôt de données aux différentes étapes du cycle de vie de la donnée ? Comment trouver et choisir un entrepôt de données ? ppt - [https://anfdonnees2017.sciencesconf.org/data/pages/Entrepots\\_ANFRenatis\\_2017\\_Cocard\\_Aventurier\\_1.pdf](https://anfdonnees2017.sciencesconf.org/data/pages/Entrepots_ANFRenatis_2017_Cocard_Aventurier_1.pdf)

Déposer/Publier dans des Plateformes logicielles de gestion de données locales - (Infrastructure de données géographiques)

Les services informatiques des laboratoires de recherche et des OSU ont pour vocation et pour mission d'intervenir dans la gestion des données d'observation acquises sur le terrain. Après les phases de collecte et de transfert de données que nous avons vues dans les précédentes étapes du cycle de vie des données, il est nécessaire “in fine” de se préoccuper de la valorisation des données en permettant la facilité d'accès et leur réutilisation.

Un certain nombre de logiciels font office de plateforme d'accès et de gestion des données.

Ils permettent de présenter les données avec leur métadonnées, de fournir des interfaces de recherche, de géolocaliser les données, et parfois de visualisation des données avec des graphes. La gestion des données demande donc aux informaticiens concernés de développer des compétences dans des domaines très divers (gestion de la production des données, organisation de la donnée, gestion de la géolocalisation, gestion de la donnée dans le temps, mise à disposition et visualisation des données, modèle OAIS, ...).

Les événements suivants organisés par le réseau SIST fournissent un certain nombre de connaissances sur l'utilisation d'infrastructure de données géographiques (IDS, IDG) et de plateforme logicielles de gestion des données

infrastructure de données spatiales et de traitements GEOSUD : des standards à la réalité

[https://sist16.sciencesconf.org/data/pages/01\\_JC\\_Desconnets.pdf](https://sist16.sciencesconf.org/data/pages/01_JC_Desconnets.pdf) - Jean-Christophe Desconnets, UMR Espace-Dev

OSUNA : Mise en place d'une IDS pour le programme de recherche Réseau de Suivi et de Surveillance de l'Environnement.

[https://sist16.sciencesconf.org/data/pages/02\\_L\\_Salaun.pdf](https://sist16.sciencesconf.org/data/pages/02_L_Salaun.pdf) - LOIC SALAUN, Observatoire des Sciences de l'Univers Nantes Atlantique

Sextant, infrastructure de données géographiques marines et littorales [https://sist16.sciencesconf.org/data/pages/03\\_M\\_Sextant.pdf](https://sist16.sciencesconf.org/data/pages/03_M_Sextant.pdf) - Catherine Satra Le Bris, Ifremer - IDM/SISMER - Mickael Treguer, Ifremer

Parc National de planeurs sous-marins

[https://sist16.sciencesconf.org/data/pages/04\\_K\\_Bernardet.pdf](https://sist16.sciencesconf.org/data/pages/04_K_Bernardet.pdf) - Karim Bernardet, Division Technique de l'INSU

GBIF - Système Mondial d'Information sur la Biodiversité

[https://sist16.sciencesconf.org/data/pages/07\\_ME\\_Lecoq.pdf](https://sist16.sciencesconf.org/data/pages/07_ME_Lecoq.pdf) - Marie-Elise Lecoq, Muséum national d'histoire naturelle

Utilisation de plateformes d'accès aux données Copier les succès et rester simple (AMEO) : Mise à disposition de sorties de modèles climatiques avec un NAS, THREDDS et ERDDAP.

[https://sist16.sciencesconf.org/data/pages/11\\_T\\_Valero\\_F\\_Bongat.pdf](https://sist16.sciencesconf.org/data/pages/11_T_Valero_F_Bongat.pdf) - Thierry VALERO, Institut de Recherche pour le Développement, Laboratoire d'Océanographie et du Climat : Expérimentations et Approches Numérique

Autres guides de bonnes pratiques:

<https://hal.archives-ouvertes.fr/hal-01275841/document> (2016 MC Jacques Y. Arnould "Guide de bonnes pratiques Gestion et valorisation des données")

<https://hal.archives-ouvertes.fr/hal-01908766> (2019 Lionel Maurel "La

<https://espace-mondial-atlas.sciencespo.fr/fr/>

