



HAL
open science

OGIIS : Étude de faisabilité d'un outil de gestion intégrée de l'information scientifique

Guillaume Fantino, Kristell Michel, Maxime Maignan, Emilie Lerigoleur

► To cite this version:

Guillaume Fantino, Kristell Michel, Maxime Maignan, Emilie Lerigoleur. OGIIS : Étude de faisabilité d'un outil de gestion intégrée de l'information scientifique. [Rapport de recherche] OHM Vallée du Rhône; Réseau des OHM; OHM Haut Vicdessos. 2013. hal-03527058

HAL Id: hal-03527058

<https://hal.science/hal-03527058>

Submitted on 14 Jan 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



ohm
Observatoires
Hommes-Milieus

OGIIS

ÉTUDE DE FAISABILITÉ D'UN OUTIL DE GESTION INTÉGRÉE DE L'INFORMATION SCIENTIFIQUE

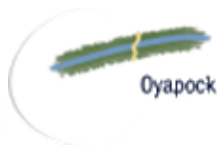
Rapport complet (document de travail)

Guillaume Fantino, UMR EVS – CNRS, OHM Vallée du Rhône

Kristell Michel, UMR EVS – ENS Lyon, OHM Vallée du Rhône

Maxime Maignan, UMR IMBE – CNRS, Réseau des OHM

Emilie Lerigoleur, UMR GEODE – CNRS, OHM Haut Vicdessos



REMERCIEMENTS

Les Observatoires Hommes-Milieus sont placés depuis leur création dans une dynamique d'interactions avec les sphères publiques originales par la modernité des questions et des pratiques qui s'y développent. En créant ces outils, aujourd'hui reconnus Laboratoire d'Excellence par le Ministère de l'Enseignement Supérieur et de la Recherche dans le cadre des Investissements d'Avenir, l'Institut Ecologie et Environnement du Centre National de la Recherche Scientifique a offert un espace et un contexte innovant pour une réflexion interdisciplinaire sur les enjeux relatifs aux données scientifiques. C'est de ce creuset que cette étude est issue et ses auteurs remercient à ce titre celles et ceux qui le font exister. Plus particulièrement, cette étude n'aurait pas eu lieu sans le concours des OHM eux-mêmes, de leurs directeurs et directrices, de leurs ingénieurs, de la direction du Dispositif de Recherche Interdisciplinaire sur les Interactions Hommes-Milieus et de Monsieur Robert Chenorkian, Directeur Adjoint Scientifique en charge de la politique de site de l'INEE.

Quand ils n'étaient que propositions, les travaux restitués dans ce document ont reçu l'appui financier des OHM Bassin minier de Provence et Vallée du Rhône. Nous sommes reconnaissants pour le crédit que nous ont accordé leurs directions. Cette confiance et ces moyens nous ont permis de réunir les conditions nécessaires aux échanges dont ce projet s'est nourri. Nous avons aussi bénéficié de l'expertise de l'Unité Mixte de Service « Bases de données sur la Biodiversité, Ecologie, Environnement et Sociétés » (UMS 3648 BBEES) dont nous remercions la directrice, Cécile Callou, et les personnels participant au projet : Chloé Martin, Amandine Hénon, Philippe Grison, Aurélien Maire et Dominique Vallée.

Dans la démarche consultative d'OGIIS, toutes leurs contributions ont été essentielles et nous percevons déjà la diversité des rencontres, des propos et des perspectives suscités comme autant de retombées positives. En souhaitant que le travail accompli fournisse un terreau fertile pour les prochaines collaborations, nous saluons spécialement la participation des membres d'OHM et des intervenants extérieurs qui, par leurs soutiens, leurs conseils, leurs avis, leurs relectures attentives ou leur attention, ont donné à ce travail toute sa dimension collective : Romain David, Damien Davy, Axel Ducourneau, Florence Dujardin, Chloé Durif, Jules Fleury, Gérald Foliot, Didier Graillet, Wilfried Heintz, Laurence Nicolas, Jean-Claude Raynal, Samuel Robert, Armand Rotereau, Antoine Zimmermann.

TABLE DES MATIERES

Introduction.....	4
Un constat généralisé dans le domaine de la recherche	4
Qu'en est-il des OHM en matière de gestion des données ?.....	5
Partie 1. La faisabilité d'un outil commun de gestion de l'information scientifique.....	7
I. Les potentialités du ROHM pour la mise en œuvre d'un outil commun	7
I.1. Des limites à l'harmonisation des pratiques et au développement d'une plate-forme commune de gestion de l'information scientifique	8
I.1.1. Des observatoires Hommes-Milieus creusets de diversités	8
I.1.2. Contraintes juridiques et données scientifiques.....	9
I.1.3. Moyens humains et temps nécessaire à la gestion de l'information scientifique	11
I.2. Les facteurs favorables à l'entreprise d'OGIIS au niveau du ROHM	11
I.2.1. Un contexte technologique et réglementaire favorable	11
I.2.2. Rationalisation de la gestion des données dans le cadre d'observatoires scientifiques	12
I.2.3. Un cadre d'action commun à tous les OHM	13
II. Présentation générale d'OGIIS : fonctionnalités et architecture.....	14
II.1. Les fonctionnalités retenues	14
II.1.1. Recensement des données existantes (PHASE 1)	15
II.1.2. Conservation des lots de données sur le long terme (PHASE 1).....	15
II.1.3. Consultation/visualisation des données (PHASE 2)	16
II.1.4. Partage des données/téléchargement (PHASE 2).....	16
II.2. L'architecture globale de la plate-forme OGIIS.....	18
II.2.1. Un hébergement unique, une administration décentralisée, un accès généralisé (en tout lieu, en tout temps).....	18
II.2.2. Une structure modulaire évolutive	19
III. Les modalités techniques et organisationnelles pour la mise en place d'OGIIS.....	20
III.1. Spécifications organisationnelles	21
III.1.1. Administrateurs et Utilisateurs	21
III.1.2. Interactions entre les différents groupes et entre les instances de décision du DRIIHM et du ROHM	22
III.2. Les spécifications techniques	24
III.2.1. Type de données et choix des normes de description.....	24
III.2.2. Choix de la (ou des) solution(s) technique(s).....	27
III.2.3. Authentification et accès à la plateforme OGIIS.....	29
III.2.4. Matériels requis et hébergement	30
III.3. Proposition de planning et recommandations pour la poursuite du projet.....	31
Partie 2. méthodes, résultats et préconisations pour la mise en place d'OGIIS	33
I. Méthodologie de travail pour le projet OGIIS.....	33
I.1. Déroulement du projet	33
I.1.1. Phase 1	33
I.1.2. Phase 2	33
I.2. Méthodes et outils	34
I.2.1. Recensement par formulaire	35
I.2.2. Échanges et réunions	35

II. Rappel des principaux résultats	35
II.1. Recensement de l'existant : outils, contraintes, attentes	36
II.2. Définition des spécifications générales au niveau de la plateforme	37
II.2.1. Hébergement	37
II.2.2. Supervision et maintenance.....	37
II.2.3. Intégration web	37
III. Spécifications techniques des différents modules d'OGIIS.....	38
III.1. Construction du référentiel acteurs/projets.....	38
III.1.1. Mise en perspective	38
III.1.2. Propositions pour une nouvelle gestion intégrée des projets de l'APR du ROHM	39
III.1.3. Détail de la modélisation de la base de données du référentiel acteurs/projets	42
III.2. Construction du référentiel métadonnées	44
III.2.1. Formulaire de saisie.....	44
III.2.2. Consultation	45
III.2.3. Gestion dynamique des métadonnées	45
III.3. Construction du module d'archivage	47
III.4. Construction du module géo-visualisation	47
III.5. Construction du module stockage/partage	48
III.5.1. Réception des données	48
III.5.2. Intégration.....	48
III.5.3. Interfaces.....	49
III.5.4. Gestion des flux.....	49
III.5.5. Hébergement	49
Conclusion & perspectives	51
Observations et informations dans le système OHM	51
OGIIS : vers un nouveau type de relations humains-OHM ?	51
Acronymes / Sigles	53
Bibliographie	54
Annexes	56
Liste indicative et non-exhaustive des correspondances entre éléments de 2 modèles de description de ressources : DC et INSPIRE	57
Formulaire d'inscription à OGIIS (vue factice)	60
Formulaire de réponse aux APR (vue factice)	61
Formulaire de sélection des réponses aux APR (vue factice)	67
Formulaire de saisie des rapports de restitution (vue factice)	68
Modèle logique de données.....	73

INTRODUCTION

OGIIS a été financé au titre de l'appel à projet (APR) 2012 du Réseau des Observatoires Hommes-Milieux¹ (ROHM). Il a également bénéficié d'un soutien financier de deux Observatoires Hommes-Milieux : l'OHM Bassin Minier de Provence (OHM-BMP) et l'OHM Vallée du Rhône (OHM-VR). Sa finalité est la réalisation d'une étude de faisabilité pour la mise en place d'Outils communs de Gestion Intégrée de l'Information Scientifique (OGIIS). Ce projet s'est construit autour d'une communauté préexistante au sein du Réseau OHM, travaillant sur la question des modalités de gestion des bases de données produites et/ou gérées par les OHM. La nécessité de ce travail a été réaffirmée dernièrement, au cours du séminaire annuel ROHM, organisé à Aveiro, le 18 juin 2013.

UN CONSTAT GENERALISE DANS LE DOMAINE DE LA RECHERCHE

Les OHM, en tant qu'observatoires scientifiques, gèrent des volumes importants de données. Comment les gérer efficacement ? Cette question dépasse largement le cadre des OHM. En effet, le réseau des ZA (Zones Ateliers, autre grand équipement de l'INEE) y est également confronté et tente aussi de trouver des solutions, individuelles ou collectives, comme le suggère le séminaire « *Métadonnées : concepts et outils* » qui s'est tenu à Rennes, le 26 mai 2011.

Conscient des enjeux forts existants autour de la gestion de la donnée scientifique et du retard accumulé par la recherche française sur cette question, l'INEE a inscrit l'année 2011 comme « l'année des bases de données ». Il a d'ores et déjà engagé un important travail de recensement des bases de données des équipes de recherche INEE (<http://www.bdd-inee.cnrs.fr/>). Les journées *Prospectives de l'INEE* (Avignon, 2012) ont annoncé la prolongation de ces efforts. La création de l'UMS 3468 BBEES (Bases de données sur la Biodiversité, Ecologie, Environnement et Sociétés), le 1^{er} septembre 2011, suivie en mai 2012 de l'émergence du RBDD (Réseau Bases De Données) – aujourd'hui rattaché à Mission Interdisciplinarité du CNRS, confirme l'implication de l'INEE et sa volonté d'aider les chercheurs français à faciliter l'insertion de leurs bases de données dans des dispositifs nationaux et internationaux.

L'INSHS (l'Institut des Sciences Humaines et Sociales) est également concerné par ce sujet. Pour l'aider à avancer sur cette thématique, il bénéficie du support technique de la Très Grande Infrastructure (TGIR) des Humanités Numériques (Huma-Num) dont les objectifs sont les suivants : i) ouvrir à la communauté un ensemble de services mutualisés à haute valeur ajoutée, ii) constituer la mémoire vivante de la recherche en SHS et donner les moyens à la communauté de l'exploiter. La TGIR Huma-Num met ainsi à disposition des équipes de recherche (SHS principalement) des outils de gestion pour certains types de données (ex: [ISIDORE](#), [MédiHal](#), [la grille ADONIS](#), [NUMES](#)).

L'accessibilité aux données est devenue un enjeu capital, tant du point de vue de la recherche que de ses interactions avec les sphères opérationnelles. En effet, celle-ci participe activement à la valorisation des travaux de recherche, tout en répondant, de plus en plus, à des obligations réglementaires. La [Directive européenne INSPIRE](#) par exemple concernent particulièrement les OHM en tant que producteurs de données. Cette directive stipule qu'en décembre 2013, les « métadonnées des données et des services relatifs aux

¹ Les Observatoires Hommes-Milieux (OHM) constituent un des grands équipements de l'Institut Environnement Ecologie (INEE) du Centre National de la Recherche Scientifique (CNRS). Ils participent à la promotion de l'interdisciplinarité pour une recherche en écologie globale. Pour plus d'information : <http://www.ohm-inee.cnrs.fr/spip.php?rubrique1>.

thèmes de l'annexe III » devront être accessibles en ligne. Parmi ces thèmes, figurent notamment les « installations de suivi environnemental ». De plus, tous (chercheurs et acteurs publics) s'accordent pour dire qu'une meilleure gestion des données scientifiques est souvent à l'origine de plus d'interdisciplinarité.

Cela étant, ces dernières années, les progrès technologiques ont permis la mise en place d'instrumentation produisant des volumes de données toujours plus importants (phénomène du « Big Data ») qui obligent notamment les équipes de recherche à réfléchir à des stratégies (de stockage, de diffusion, de sauvegarde et de traitement) de plus en plus élaborées pour la gestion de leurs données.

QU'EN EST-IL DES OHM EN MATIERE DE GESTION DES DONNEES ?

Très tôt, le collectif des OHM s'est mobilisé autour de la donnée : le projet OGIIS est la continuité logique d'une réflexion engagée dès 2011, au sein du groupe de travail « base de données » du ROHM et qui vise à définir les synergies nécessaires au développement d'un « système d'information durable » : c'est-à-dire une organisation des flux de données qui assume sur le long terme son rôle dans la maîtrise de la connaissance².

La mise en place d'une infrastructure technique pour la recherche peut parfois aboutir à un résultat non conforme aux objectifs initiaux car le processus de définition des besoins, la réalisation technique et la mise en production repose sur des éléments de compréhension implicite entre des personnes d'un même collectif (ici, le ROHM), mais dont les fonctions et les compétences diffèrent (informaticiens vs. thématiciens, ingénieurs vs. chercheurs/directeurs). C'est pourquoi, le groupe « base de données » du ROHM a souhaité réfléchir, via le projet OGIIS, à la faisabilité de créer une plate-forme de gestion des données à caractère scientifique en incluant cette diversité des points de vue et des compétences, intrinsèque aux OHM.

OGIIS doit ainsi permettre de définir les conditions matérielles, techniques et organisationnelles à mettre en œuvre pour construire un *système d'information durable* pour le ROHM, qui a pris, en mars 2012, la forme d'un Laboratoire d'Excellence (LabEx) ayant pour nom le DRIIHM (Dispositif de Recherche Interdisciplinaire des Interactions Hommes-Milieus).

Il est cependant important de noter que ce projet s'est voulu dès le début plus englobant en souhaitant bénéficier des expériences menées par d'autres équipes et en les incluant au mieux dans la démarche de co-construction du projet. En favorisant cette dynamique collective de réflexion et de production, OGIIS a exprimé la volonté d'éviter quelques écueils parmi ceux qui peuvent survenir dans la mise en place d'importantes infrastructures numériques, dans un contexte de demande accrue de moyens et de visibilité, à savoir :

- des divergences entre concepteurs, réalisateurs et utilisateurs,
- un manque d'évolution de l'outil en lien avec les besoins régulièrement exprimés,
- une énième initiative exigeant, à nouveau, d'importantes ressources matérielles,
- une non-adhésion au dispositif et une moins-value en termes de retombées scientifiques.

Prochainement, le DRIIHM envisage le recrutement d'un ingénieur d'études dédié aux travaux sur les bases de données. Les réflexions menées dans le cadre du projet OGIIS doivent aider à préciser les besoins des équipes en matière de gestion de données. Indirectement, elles précisent les compétences et actions qui

2 « [...] Ce n'est pas l'information elle-même qui est au cœur de la révolution technologique, c'est la capacité à accéder à tout moment et en tout lieu à l'information pertinente, à créer une information innovante et à la partager [...] Les limites de la croissance liées au tarissement des ressources naturelles accélèrent l'émergence d'une nouvelle économie fondée sur l'immatériel, non polluante et orientée vers le développement durable. Elle exploite la connaissance comme matière première en vue de favoriser une croissance moins dépendante des seuls biens matériels. [...] » [Tassin 2010, pp. 209- 246].

seront requises pour répondre à leurs attentes ; elles posent donc un cadre dont cette activité pourra bénéficier.

PARTIE 1. LA FAISABILITE D'UN OUTIL COMMUN DE GESTION DE L'INFORMATION SCIENTIFIQUE

L'ensemble des OHM forment un réseau (ROHM) qui sert d'espace de discussions et d'échanges autour de thématiques partagées. C'est à l'occasion du 1^{er} séminaire du ROHM (Aix-en-Provence, juin 2010) qu'a été abordée, pour la première fois, la question des systèmes d'information (SI) au sein des OHM. Peu de temps après, un groupe de travail « base de données » s'est constitué. Il visait à échanger sur les pratiques en cours en matière de gestion des données scientifiques collectées et/ou produites dans les différents OHM. Lors du 2^e séminaire du ROHM (Toulouse, juin 2011), un principe a été adopté : les OHM doivent se munir d'un SI permettant à minima un recensement des données produites. Ces SI peuvent reposer sur des outils et/ou technologies divers et variés, mais le ROHM veillera à assurer l'interopérabilité des solutions techniques retenues par chacun des OHM.

Cette règle de conduite présente un inconvénient majeur : celui de permettre le développement de SI à deux vitesses. Les OHM dotés de moyens humains compétents en la matière et bénéficiant de surcroît de solutions techniques d'ores et déjà fonctionnelles sont inéluctablement avantagés. Même si le ROHM a pour mission d'assurer le partage des expériences et un soutien technique et matériel aux OHM les moins avancés, la question de la mutualisation des outils a rapidement émergé. Le groupe « bases de données » a initié le projet OGIIIS pour apporter les premiers éléments de réponse à cette interrogation.

Cette première partie est une synthèse des réflexions et des préconisations élaborées dans le cadre du projet OGIIIS. Dans un premier temps, elle aborde le contexte dans lequel est né OGIIIS et les besoins exprimés au sein du ROHM. Dans un second temps, elle contribue à la définition de ce que pourrait être un outil de gestion de l'information scientifique, commun à tous les OHM.

I. LES POTENTIALITES DU ROHM POUR LA MISE EN ŒUVRE D'UN OUTIL COMMUN

Pour aider à la construction d'OGIIIS, plusieurs sessions d'échange ont été organisées (tenue de réunions et soumission de questionnaires). Elles ont permis d'établir un état de l'art des outils de gestion de données existant dans chaque OHM. Ce travail a pu se faire grâce à l'identification et à la participation active et volontaire des ingénieurs, présents dans les observatoires. Ces rencontres ont également aidé à la définition des points de blocage, réels ou potentiels. Enfin, les réseaux d'échanges privilégiés établis par chaque OHM ont permis d'introduire dans la construction d'OGIIIS des chercheurs et des ingénieurs d'autres structures, compétents sur ces questions d'administration de la donnée, et d'apporter ainsi une expertise extérieure nécessaire à l'épanouissement et à l'acceptation de ce projet par le plus grand nombre. En effet, ils ont su fournir les arguments pour (ré-)affirmer l'intérêt de disposer d'une plateforme collaborative de gestion de l'information au niveau d'un collectif ayant un sens et des moyens pour la co-construire dans un contexte global, favorable à ce type d'initiative au niveau de la recherche.

I.1. DES LIMITES A L'HARMONISATION DES PRATIQUES ET AU DEVELOPPEMENT D'UNE PLATE-FORME COMMUNE DE GESTION DE L'INFORMATION SCIENTIFIQUE

La première limite identifiée par le réseau pour la mise en place d'une plateforme de gestion et de valorisation de l'information scientifique est en lien étroit avec les spécificités de chaque OHM et les pratiques qui en découlent. La composante juridique est, par ailleurs, un élément à ne pas négliger : les données relèvent toutes du « droit d'auteur ». Mais que ce signifie précisément cette expression ? Qu'en est-il des données à caractère scientifique produites par des agents publics ? Enfin, tout projet informatique nécessite du temps, des moyens humains et matériels et la définition des missions de chacun. Il n'est pas possible d'ignorer ces contraintes, sous peine de penser et de développer un outil non adapté.

I.1.1. DES OBSERVATOIRES HOMMES-MILIEUX CREUSETS DE DIVERSITES

La diversité des OHM s'exprime au moins à 3 niveaux différents, présentés et explicités ci-après :

- **Diversité thématique** : si tous les OHM traitent des interactions hommes-milieux, les territoires étudiés impliquent l'exploration de champs thématiques qui ne font pas appel aux mêmes conjonctions de disciplines. Ainsi, les OHM s'intéressant à des territoires urbanisés (le bassin minier de Provence, Estarreja, le littoral méditerranéen) sont plus facilement en prise avec les questions de pollutions industrielles et leurs impacts sur les populations. Dès lors, leurs liens avec des disciplines connexes, telles que la géochimie, la toxicologie, la sociologie, l'épidémiologie, etc. se font naturellement. D'autres territoires étudient les écosystèmes forestiers (Haut-Videssos) ou fluviaux (Vallée du Rhône). Les OHM internationaux (OHM.I), quant à eux, impliquent plus fréquemment des disciplines comme l'ethnologie ou l'anthropologie. Au-delà des différences de territoires, la diversité thématique est intrinsèque à la nature même des OHM, dont l'un des objectifs principaux est de croiser les regards et de produire des approches scientifiques pluridisciplinaires. Cette caractéristique implique la gestion de lots de données très différents du point de vue de leur structure, de leur taille et par voie de conséquence, de leur format numérique.

- **Diversité technique (outils et compétences)** : cette diversité de formats oblige les OHM à s'intéresser à des outils de gestion des données variés. En effet, la plupart du temps, la construction de ces outils répond à des besoins spécifiques, exprimés par une communauté d'utilisateurs spécialistes. Ainsi, les équipes travaillant avec des données spatialisées (ou données géographiques) développent et utilisent des solutions techniques où la géolocalisation, la mention de références spatiales normalisées et la visualisation cartographique ont une place centrale. C'est le cas des OHM BMP³ et VR qui ont développé leurs propres géocatalogues⁴ (ou géorépertoires) basé sur l'application Open Source [GeoNetwork](#) (GN). Au lancement de l'étude OGII, peu d'OHM disposaient d'outils spécifiques pour gérer leurs données et leurs métadonnées. La seule convergence existante reposait sur la gestion des données spatiales avec un couplage **GeoNetwork** (catalogue de métadonnées⁵)/**ArcGis Server** (visualisation de services web cartographiques).

³ Les OHM sont actuellement au nombre de huit. Nous utiliserons pour les besoins de cette étude les sigles suivant : **BMP** pour Bassin Minier de Provence ; **OYA** pour Oyapock ; **HV** pour Haut-Videssos ; **EST** pour Estarreja ; **GMV** pour Tessékéré ; **VR** pour Vallée du Rhône ; **LM** pour Littoral méditerranéen ; **NUN** pour Nunavik.

⁴ Un géocatalogue permet d'organiser des métadonnées de données à référence spatiale, structurées de façon identique afin de les valoriser et d'en favoriser la diffusion et la consultation pour finalement faciliter le partage et la mutualisation de données géographiques.

⁵ Catalogue de métadonnées de l'OHM BMP : <http://cat-bmp.ohm-provence.org/geonetwork/srv/fr/main.home> et de l'OHM VR : <https://www.ohm-provence.org/spip.php?rubrique18>. Visualisation de *webservices* cartographiques pour l'OHM BMP : <https://www.ohm-provence.org/spip.php?rubrique18> et pour l'OHM VR :

La diversité des outils est en lien direct avec le personnel dont dispose chaque OHM (nombre d'ingénieurs et de techniciens, permanents vs. contractuels, compétences et spécialités). Ainsi, la mise en place d'outils de gestion des données numériques exige, le plus souvent, des connaissances en informatique (gestion de serveurs, gestion de bases de données) qui ne sont que rarement présentes dans les OHM. Certains développements peuvent donc être freinés du fait de cette absence de personnel qualifié. Mais une fois réunies au niveau du ROHM, les compétences variées des personnels des OHM offrent l'avantage de la complémentarité.

▪ **Diversité géographique (situation et partenaires privilégiés)** : la situation géographique est une autre source de diversité entre OHM. Certains sont métropolitains (BMP, HV, VR, LM) d'autres sont présents dans les DROM (Département et Région d'outre-mer) tel que l'OHM OYA, les derniers se situent à l'étranger dans des régions fortement industrialisées (EST), susceptibles de l'être (NUN) ou encore essentiellement rurales (GMV). Ces trois derniers observatoires constituent les OHMI.

Le contexte doit être considéré pour la mise en place d'outils de gestion des données/métadonnées. Cet élément détermine :

- le niveau d'infrastructure technologique (matériel informatique et connexion/débit internet),
- la qualité et la quantité des données numériques formalisées,
- le statut des données qui peuvent appartenir à des chercheurs ou organismes non français.

Ces éléments forment un champ de contraintes à prendre nécessairement en compte pour la mise en place d'un outil commun. L'aspect partenarial renvoie, d'une manière plus large, à la question de la propriété des données qui s'impose dans le domaine de la gestion collective des données.

I.1.2. CONTRAINTES JURIDIQUES ET DONNEES SCIENTIFIQUES

La question du statut juridique des données est complexe et souvent, le versement de données dans des outils partagés soulève la question de la propriété des données. En effet, les données scientifiques produites par des *chercheurs* sont soumises au droit d'auteur qui recouvre des droits moraux et patrimoniaux. En revanche, cela n'est pas le cas pour les *ingénieurs* travaillant sur des travaux de recherche : leur établissement de rattachement est le propriétaire de la donnée produite, notamment dans les cas où il est établi que la méthode et la commande correspondent à une exigence hiérarchique supérieure. De ce point de vue, il n'y a pas, pour les chercheurs producteurs de données, d'obligations juridiques à fournir leurs lots de données, dans la mesure où cela relève de leurs droits moraux.

Les chercheurs doivent cependant respecter l'engagement que constitue la signature de la Charte des OHM, qui finance tout ou partie de leurs études. Cette dernière stipule en effet que : « *les données et résultats issus des recherches menées dans le cadre [d'un] OHM appartiennent aux établissements dont relèvent les partenaires. Cependant, les partenaires reconnaissent au projet de l'OHM le droit de constituer un système d'information. Ils acceptent le principe de confier tout ou partie de leurs résultats à la base de données de l'OHM, accessible à tous les partenaires selon des droits d'accès qui seront définis au cas par cas* »⁶. Les lots de données transférables aux OHM sont donc à la discrétion des chercheurs, avec néanmoins l'obligation de fournir au moins une partie des résultats.

<http://ohm-vr.org/sig.php>.

⁶ Charte des OHM disponible en ligne sur le site du ROHM : <http://www.ohm-inee.cnrs.fr/spip.php?article196>

En fait, la question de la mise à disposition des résultats de recherche se pose véritablement pour les chercheurs impliqués dans un projet OHM, mais non directement (personnellement) ou indirectement (son équipe de rattachement) signataires de la charte. La question de la propriété des données et de la possibilité de leur diffusion se pose aussi dans le cas des données produites par des financements composites (un ou des OHM et un autre contrat de recherche). Cette question dépassant le cadre des OHM, ceux-ci peuvent solliciter des instances compétentes en la matière : groupe de travail "Aspect Juridiques" du Réseau Bases de Données (RBDD), services Partenariats et Valorisation (SPV) des Délégations Régionales (DR), voire la Direction des Affaires Juridiques (DJA) du CNRS. Il n'en resterait pas moins aux OHM de mettre en place des relations de confiance et de bonne volonté propices au partage de l'information, de pratiquer la transparence quant aux exigences du dispositif OHM en amont de la livraison des projets et d'élaborer, le cas échéant, des mécanismes incitatifs efficaces.

Les enjeux relatifs à la propriété des données sont réels car l'information scientifique requiert souvent un temps long de collecte ou de production. Les objets ainsi créés matérialisent un savoir-faire valorisable, parfois financièrement, mais le plus souvent scientifiquement via des publications. Certains chercheurs sont très sensibles à cette nécessaire protection de la propriété de leurs données tant que les publications n'ont pas été produites.

Sur le plan juridique, une autre limite à la mise en place d'une plate-forme collective de gestion des données est celle induite par les données nominatives (i.e. contenant des informations personnelles : nom, prénom, adresse, sexe, âge...). Toute diffusion informatique de données personnelles est soumise à une déclaration préalable à la Commission Nationale de l'Informatique et des Libertés (CNIL) afin de s'assurer : 1) que les mesures de sécurité physiques (sécurité des locaux) et logiques (sécurité des systèmes et réseaux) ont été mises en place et qu'elles sont adaptées à la nature des données et aux risques présentées par leur traitement, 2) que seules les personnes autorisées peuvent accéder aux données personnelles contenues dans un fichier, 3) que les données personnelles ont une date de péremption⁷, 4) que les personnes concernées par un fichier ont bien été informées de leurs droits⁸. Il est donc impératif de veiller à préserver l'anonymat, en particulier pour certains jeux de données sensibles (enquêtes et entretiens). Sur ce point, il est possible de contacter un référent CNIL, présent au sein du CNRS et dans chaque université ou établissement rattachés au Ministère de l'Enseignement Supérieur et de la Recherche qui disposent d'un service informatique conséquent.

⁷ Les travaux de la commission européenne à ce sujet sont à suivre ; voir notamment [COMMISSION EUROPEENNE 2012]

⁸ <http://www.cnil.fr/>

I.1.3. MOYENS HUMAINS ET TEMPS NECESSAIRE A LA GESTION DE L'INFORMATION SCIENTIFIQUE

La contrainte temporelle est un facteur à bien considérer. Le temps nécessaire au développement d'une infrastructure de recherche peut s'avérer long, mais il est généralement bien défini préalablement au lancement d'un projet. En revanche, le temps nécessaire à la maintenance et à l'implémentation d'une infrastructure est plus difficile à estimer. Il s'agit souvent d'une administration légère au quotidien, mais qui exige une vigilance régulière et des compétences « expertes ». ***Si pour le développement ou l'alimentation d'un outil, l'emploi de personnel temporaire est envisageable, il est préférable que l'administration et la maintenance soient gérées par du personnel permanent*** pour être assuré du bon fonctionnement de l'application et de l'accès à ses fonctionnalités sur le long terme – sous une forme pouvant évoluer au cours du temps.

La question des moyens humains mis à disposition renvoie à celles des compétences (cf. 1.1.1.1. Des Observatoires Hommes-Milieus, creusets de diversités). Les différentes tâches, mises en œuvre pour la création et l'utilisation d'une plate-forme telle qu'OGIIS, impliquent des compétences diverses, disponibles dans les OHM, mais qui dépassent le seul corps des ingénieurs/techniciens. En effet, il est nécessaire de faire la distinction entre a) la création des outils, b) leur administration courante et c) leur exploitation incluant notamment la phase de saisie des métadonnées. Dans les deux premiers cas (a et b), disposer de compétences techniques s'avère nécessaire – même si, le plus souvent, il est possible de dissocier le travail de mise en place de l'outil qui exige un haut niveau de technicité informatique et le travail d'administration au quotidien que tout un chacun peut apprendre à maîtriser. Dans le troisième cas (c), peu de compétences techniques sont *a priori* nécessaires. En revanche, il faut disposer de connaissances sur l'ensemble des thématiques couvertes par un OHM, pluridisciplinaire par nature. Par ailleurs, les différents retours d'expériences aboutissent tous à la même conclusion : ***les producteurs de données (i.e. les scientifiques) sont souvent les personnes les plus à même de décrire la donnée***. Lorsque les tâches d'administration et d'implémentations incombent à une seule et unique personne (en plus des missions qui lui sont confiées dans le cadre de ses fonctions), ce travail, jugé « *ingrat* » par beaucoup parce que répétitif et peu stimulant, peut vite accaparer tout son temps. Ce constat explique en partie le retard accumulé en matière d'accès à la donnée au sein de la recherche.

Aujourd'hui, le réseau (ROHM) dispose d'ingénieurs référents au sein de la presque totalité des OHM. Le DRIIHM, en tant que LabEx, est une source de moyens supplémentaires qui devraient permettre de pallier les manques techniques et humains, nécessaires à la mise en place d'une infrastructure comme OGIIS. Et si la diversité du réseau peut générer des contraintes fortes pour le développement d'une plate-forme commune, ce même réseau est aussi un atout pour le projet OGIIS.

I.2. LES FACTEURS FAVORABLES A L'ENTREPRISE D'OGIIS AU NIVEAU DU ROHM

Les enjeux forts autour de la gestion de la donnée et les perspectives favorables qu'apporte le LabEx DRIIHM font du ROHM un lieu propice aux échanges et à l'élaboration d'une plate-forme commune de gestion de l'information scientifique. Les points clés qui font d'OGIIS un projet d'intérêt pour le collectif sont de trois ordres : 1) un contexte global propice, 2) une « rationalisation » des pratiques affirmée au niveau du ROHM et 3) une organisation et un mode de fonctionnement des OHM très similaires.

I.2.1. UN CONTEXTE TECHNOLOGIQUE ET REGLEMENTAIRE FAVORABLE

La gestion et la préservation de l'information numérique est un problème complexe dont dépendent de nombreuses structures, publiques ou privées. Les OHM, le ROHM et maintenant le DRIIHM n'y font pas exception. Le phénomène du « **Big Data** » en est la principale cause. Il correspond à l'explosion des volumes de

données numériques, liée notamment aux progrès techniques accomplis ces dernières années en matière d'acquisition (multiplication des capteurs toujours plus performants, possibilité de mesures en continu), d'échanges et de stockage de données aux formats variés (données brutes ou structurées, disponibles sous différentes formes : image, texte...). Pour faire face à cette croissance de la volumétrie des données, les scientifiques doivent se munir d'outils performants pour gérer, traiter et archiver les données. Parallèlement, les données numériques sont de plus en plus accessibles et ouvertes (phénomènes de « *l'Open Data* » et de « *l'innovation participative* »). La recherche ne peut pas rester en marge de cette évolution et se retrouve aujourd'hui confrontée à l'ouverture de ses archives numériques. Cette évolution est, pour partie, encadrée réglementairement par une directive européenne, la directive INSPIRE⁹, qui vise à établir une infrastructure d'information géographique¹⁰ en Europe pour favoriser la protection de l'environnement.

Le très fort développement des technologies du web (Application Program Interface (API), Web Services (WS)) permet à présent de gérer, à distance et de manière collaborative, et/ou de visualiser des jeux d'informations numériques. Ainsi, les technologies de l'information et de la communication dans leurs évolutions les plus récentes permettent de manipuler des objets nouveaux. C'est le cas de la réalité augmentée d'objets ou de phénomènes discrets insoupçonnés du public - voire hypothétiques - et dont l'existence peut ainsi être révélée ou suggérée. Les outils numériques sont sollicités pour l'apprentissage et la formation (machine learning, le e-learning). Les langages qu'ils utilisent gagnent en évolutivité et en intelligibilité notamment grâce aux travaux et recommandations produits dans le cadre du Web sémantique. Les plateformes de partage de données offrent également la possibilité à tous citoyens d'être producteurs de données dans le cadre d'une démarche participative. Ces nouveaux outils et ces nouvelles technologies sont de plus en plus présents dans l'espace public. Il devient incontournable d'en acquérir une maîtrise suffisante pour occuper ces terrains virtuels en raison de la plus-value qu'ils apportent, notamment pour l'étude des interactions société/environnement.

1.2.2. RATIONALISATION DE LA GESTION DES DONNEES DANS LE CADRE D'OBSERVATOIRES SCIENTIFIQUES

- **Optimisation des tâches relatives à l'administration courante des données** : l'utilisation d'outils dédiés permet de simplifier, de normaliser et d'automatiser les manipulations courantes réalisées sur des jeux de données. La simplification (pré-saisie de champs à l'aide de thésaurus¹¹ par exemple) et l'automatisation (saisie automatique à l'aide de passerelles entre normes ou suggestions de saisie selon les opérations appelées par l'agent) constituent un gain de temps dans les traitements courants. Elles évitent les redondances et les incohérences liées à la diversité des opérateurs. Elles permettent de respecter plus facilement les standards partagés qui constituent la condition *sine qua none* pour une interopérabilité entre systèmes.

La rationalisation des pratiques est à développer aussi en ce qui concerne la sauvegarde et l'archivage. Aujourd'hui encore, il est possible de constater un réel déficit dans la capitalisation des connaissances en raison de pertes de données induites par des systèmes des sauvegardes trop rustiques (disques durs externes et personnels, panne matérielle sur des systèmes non dupliqués, etc.).

Cela étant, il est nécessaire de réfléchir dès à présent à des systèmes efficaces, de trouver un compromis entre la formalisation des bonnes pratiques et la préservation des marges de libertés dans la mise en œuvre de

⁹ <http://inspire.jrc.ec.europa.eu/> et http://georezo.net/blog/inspire/files/2011/03/Inspire_neophytes_V2.pdf

¹⁰ « Est entendu par infrastructure d'information spatiale un ensemble de services disponibles sur Internet, sur les sites web de différents acteurs concernés, et permettant la diffusion et le partage de données géographiques (source : *La Directive Inspire pour les néophytes*, février 2011).

¹¹ Un thésaurus est une liste organisée de termes contrôlés et normalisés représentant les concepts d'un domaine de connaissance (source : <http://fr.wikipedia.org/wiki/Th%C3%A9saurus>)

ces recommandations que ce soit en matière de sauvegarde/d'archivage ou plus largement en termes de gestion de la donnée. La rationalisation à l'échelle nationale, voire au-delà, s'initie progressivement. Mais elle ne doit pas conduire à la « standardisation » des pratiques. Elle doit préserver l'autonomie des équipes et autoriser des marges de manœuvre. Les OHM demandent un système cohérent, fiable et simple d'accès pour mieux gérer leurs données. OGIIS doit répondre à leurs attentes, mais ce projet ne doit pas gêner les initiatives individuelles. Bien au contraire, il doit contribuer à leurs développements et pouvoir poursuivre sa propre évolution c'est-à-dire son adaptation à la demande sur le long terme. C'est l'intérêt de disposer d'un collectif tel que le ROHM autour de ce projet : OGIIS alimente les observatoires qui eux-mêmes viennent nourrir à leur tour le projet grâce à leurs travaux.

- **Valorisation des données et des résultats de recherche produits dans les OHM** : la démarche scientifique repose sur une entreprise collective. L'échange des résultats est le principal moteur des avancées scientifiques. Les outils ayant pour finalité le recensement et la diffusion des données pour un meilleur partage de l'information, participent directement à cette dynamique.

Les outils de type « catalogage » facilitent la recherche de données existantes pouvant être mobilisées dans le cadre de nouveaux projets. Leur bonne exploitation limite la redondance des données et réduit de surcroît les surcoûts qu'elle occasionne. Ces outils facilitent, par ailleurs, la mise en place d'observations en continu, en portant à la connaissance des opérateurs, les secteurs déjà étudiés, les protocoles reproductibles et autres éléments utiles à l'élaboration d'un plan d'échantillonnage systématique, par exemple. Disposer d'une information détaillée et mise à jour régulièrement sur les données existantes est essentiel, notamment pour les jeunes chercheurs, doctorants et stagiaires qui ne connaissent pas l'historique de la donnée produite sur un thème ou sur un territoire donné. Cet aspect est fondamental pour des observatoires où prennent place des logiques de suivi.

Les « catalogues » mais également les outils de consultation (outils de « visualisation ») et de diffusion (outils de « partage ») permettent de donner une visibilité aux résultats de recherche. Dans la mesure où les OHM ont pour ambition d'être des observatoires de recherche à l'interface avec la sphère opérationnelle des décideurs publics, la communication sur les avancées scientifiques est primordiale. Cette valorisation vers l'extérieur favorise les levées de fond pour des projets plus ambitieux et impulsés par les OHM.

Si la visibilité sur l'extérieur est un enjeu fort pour les OHM, la connaissance de l'existant et le partage des données au sein des OHM est une nécessité (Cf. partie II), un objectif à atteindre à échéance rapide. A ce titre, la pertinence de l'approche spatiale, qui permet notamment l'intégration du temps via des approches diachroniques et l'analyse des transferts d'échelle, est rappelée ici. Au cours du dernier séminaire du ROHM (Aveiro, juin 2013), le besoin de rassembler des lots d'information spatiale communs à tous les OHM a été évoqué ; il permettrait d'engager des comparaisons entre les différents sites étudiés (approche synchronique). La base de données [Air-BMP](#)¹², développée dans le cadre de l'OHM BMP pour le suivi des particules atmosphériques, est un exemple de réflexion utile à la constitution d'un socle commun de conteneur pour les données environnementales issues de capteurs ou d'analyses, discrètes ou continues. En effet, la méthode de conception et d'échanges entre informaticien et chercheur qui en est à l'origine a abouti à un schéma logique qui, s'il n'est pas directement transposable à tous les cas, propose une organisation d'entités récurrentes dans un projet de recherche : l'étude elle-même, les opérateurs, les appareils, les sites, les mesures et les éventuels commentaires ou photos qui y sont associés, etc...

I.2.3. UN CADRE D'ACTION COMMUN A TOUS LES OHM

¹² Contacts : Yves Noack (noack@cerege.fr), Armand Rotereau (rotereau@eccorev.fr) et Jean-Claude Raynal (raynal@eccorev.fr)

Les OHM, bien que diversifiés thématiquement et géographiquement, possèdent un cadre d'action et de gestion commun. La structuration des OHM, notamment administrative et organisationnelle, est similaire - avec cependant quelques spécificités pour les OHMI. Ainsi, les OHM sont toujours adossés à une unité de recherche du CNRS qui assume notamment sa gestion financière. De la même manière, tous les OHM organisent annuellement un Appel à Proposition de Recherche (APR) qui permet de financer divers projets. Pour les projets retenus au titre d'un APR, un chef de projet est désigné et une équipe de recherche est retenue pour être le porteur financier de ce projet. Cette structure identique est un élément à considérer dans la perspective de création d'un outil commun. En effet, elle facilite un référencement homogène des lots d'information, des personnes et des projets.

II. PRESENTATION GENERALE D'OGIIS : FONCTIONNALITES ET ARCHITECTURE

Les différents éléments (atouts et contraintes), identifiés au cours des réunions de travail, ont permis d'établir un premier bilan de la faisabilité d'une plateforme commune de gestion de l'information scientifique pour le ROHM. Les limites énoncées ci-dessus ont toutefois obligé à faire des choix. Ils ont porté sur les objectifs finaux de l'outil (fonctionnalités et architecture globale – cf. section 1.II.2), sur les options techniques et organisationnelles à retenir, ainsi que sur les modes opératoires pour sa mise en place (planning) (cf. III).

II.1. LES FONCTIONNALITES RETENUES

Les fonctionnalités d'une plateforme de gestion des données sont celles du cycle de vie des données numériques, représentés ci-dessous [MICHENER 2012].

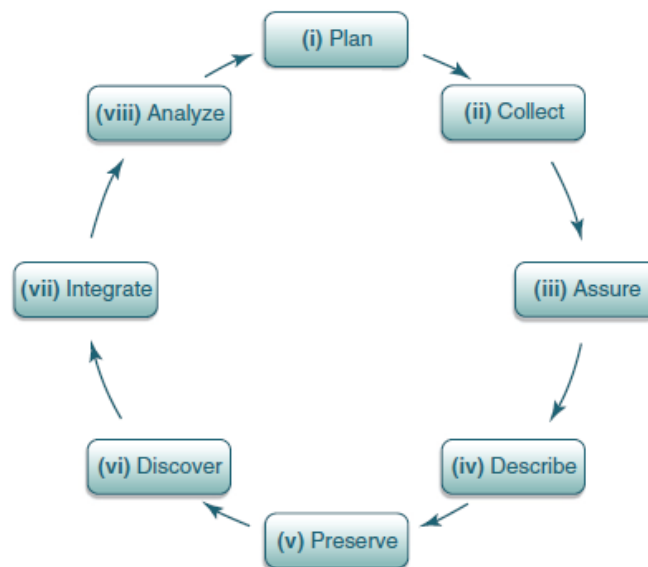


FIGURE 1 - CYCLE DE VIE D'UNE DONNEE NUMERIQUE SCIENTIFIQUE D'APRES MICHENER ET AL., 2012

La figure 1 montre les différentes étapes que la donnée est susceptible de franchir pour aboutir à la production de nouvelles connaissances : (i) planifier (e.g. un échantillonnage, une acquisition, une série d'entretiens), (ii) collecter (e.g. mesurer, interroger, observer, enregistrer), (iii) valider (i.e. confirmer la bonne qualité du jeu de données produit), (iv) décrire (i.e. cataloguer, résumer, porter à connaissance), (v) préserver (i.e. stocker, sauvegarder, archiver), (vi) découvrir (i.e. être consulté dans un cadre pertinent), (vii) intégrer (i.e. garantir l'interopérabilité sur les plans méthodologique, sémantique et logiciels), (viii) analyser (i.e. appliquer un procédé analytique formel).

Dans un premier temps, les fonctionnalités souhaitées d’OGIIS ont été identifiées, puis hiérarchisées afin de définir un ordre prioritaire d’actions quant à l’élaboration technique de l’outil à court et à moyen termes. Cette hiérarchisation, présentée ci-après, s’est formalisée lors de la phase de recensement des besoins et en lien avec les contraintes auxquelles est confronté le ROHM :

- catalogage de l’existant,
 - archivage,
 - visualisation des données,
 - stockage « organisé » de la donnée et diffusion de cette dernière.
-
- PHASE 1
- PHASE 2

Seules les deux premières fonctionnalités ont été retenues pour la première phase de développement d’OGIIS. Les deux dernières sont à planifier (cf. section 1.III.3) dès à présent en termes d’échéances et de moyens à mobiliser.

II.1.1. RECENSEMENT DES DONNEES EXISTANTES (PHASE 1)

Pour répondre à la fois aux besoins de valorisation des données existantes au sein du réseau, produites via les APR (cf. sous-section 1.I.2.2. - § valorisation des données et des résultats de recherche), la plateforme OGIIS se doit d’offrir les services d’un catalogue de métadonnées. Les catalogues de métadonnées sont des outils permettant la consultation de métadonnées¹³, c’est-à-dire des éléments descriptifs de lots de données. Ils n’offrent pas automatiquement un accès aux données. Ceci peut donc permettre, dans un premier temps, de s’affranchir des problèmes de droits afférents aux données. Un tel outil peut reposer sur une solution logicielle existante ou sur la construction d’une base de données (cf. sous-section 2.III.2.2.).

Du fait du cadre organisationnel commun des OHM, le collectif OGIIS a souhaité que le catalogue de métadonnées repose en partie sur un « référentiel Acteurs/Projets » servant à la fois de soubassement au catalogue mais aussi d’outil d’administration et de gestion des projets retenus dans le cadre des APR annuels.

II.1.2. CONSERVATION DES LOTS DE DONNEES SUR LE LONG TERME (PHASE 1)

La conservation sur le long terme des données scientifiques n’est que rarement assurée pour le moment. Les supports pour « archiver » ses données sont, pour l’heure, des plus hétérogènes en termes de supports exploités (disques durs externes, DVD, serveurs, espaces virtuels), de formats de la donnée stockée, de maintenance, de garanties logicielle et matérielle, de performance (taux de transferts, double sauvegarde...) et de localisation (gestion centralisée ou individuelle). En fait, ***l’archivage est très souvent confondu avec le stockage ou la sauvegarde.*** Pourtant ces termes regroupent des réalités différentes.

L’archivage est « *l’ensemble des actions, outils et méthodes mis en œuvre pour réunir, identifier, sélectionner, classer et conserver des contenus numériques, sur un support sécurisé, dans le but de les exploiter et de les rendre accessibles dans le temps, que ce soit à titre de preuve ou à titre informatif. Le contenu archivé est considéré comme figé et ne peut donc être modifié. La durée de l’archivage est fonction de la valeur du contenu et porte le plus souvent sur du moyen ou long terme (source : [Wikipedia](#)).* » Contrairement à une idée répandue, le stockage sur des supports de type disque dur ne garantit pas une pérennité des informations. La

¹³ Ensemble structuré d’informations servant à décrire une ressource. Les métadonnées peuvent permettre d’organiser une collection de ressources (inventaire interne) et de découvrir des ressources (moteur de recherche). Dans le cadre de recherches, les métadonnées associées à une ressource peuvent aussi permettre aux utilisateurs de mieux juger de la pertinence d’un jeu de donnée ou de sa qualité.

longévité des données enregistrées sur ce type de support numériques n'excède pas une durée moyenne de 5 à 10 ans.

OGIS doit donc permettre au ROHM de disposer d'un espace de conservation des données : il doit fournir un ensemble de moyens pour *stocker, sécuriser, pérenniser, restituer, tracer, transférer voire détruire les contenus électroniques archivés*. Ceci exige d'avoir une réflexion en amont pour choisir les formats de données adéquats (nécessairement publiés et libres) et valide, d'une part, l'importance de disposer de données descriptives normalisées, indispensables pour faire entrer les données dans un système d'archivage et d'autre part, de s'interroger sur les droits associés aux données : qui peut les consulter librement ? Existe-t-il des restrictions ? Quelles durées ont-elles ?... Cette dernière question est essentielle puisque l'archivage contiendra notamment les données de droit réservé jusqu'à leur basculement dans le domaine public, soit 70 ans après la mort du dernier ayant droit selon la législation française.¹⁴

Cet archivage offrira la possibilité de remobiliser des ressources sur un temps long, supérieure à la durée de vie du DRIHM et des OHM eux-mêmes (>10 ans).

II.1.3. CONSULTATION/VISUALISATION DES DONNEES (PHASE 2)

Ces outils reposent sur des solutions logicielles diverses car il existe bien souvent un outil conçu pour gérer un type de format particulier de données. Par exemple, une photothèque permet de consulter des données iconographiques, mais elle n'offre pas nécessairement une instance de visualisation adaptée pour de l'information spatialisée telle que des chroniques de mesures.

Beaucoup de ces outils sont des outils Open Source. Cependant, il n'est pas rare que la communauté scientifique privilégie l'utilisation ou le développement de solutions commerciales dédiés à certains formats, et qui présentent des fonctionnalités plus larges et plus abouties que les interfaces proposées sur internet.

Il est bon de noter que les quelques initiatives menées dans certains OHM pour la consultation ou la visualisation de leurs données, qu'elles soient spatiales, iconographiques, relevant de la littérature grise ou recensée répondent toutes au besoin de visibilité de ces derniers.

II.1.4. PARTAGE DES DONNEES/TELECHARGEMENT (PHASE 2)

Cette fonctionnalité, comme la précédente (« visualisation des données »), est prévue dans un second temps de développement, postérieur au catalogage de la donnée. Elle représentera un aboutissement majeur au regard de l'objectif initial d'OGIS.

L'accès aux données nécessite une infrastructure matérielle, à savoir un espace de stockage. L'accès peut se faire par FTP (File Transfert Protocole) via une interface telle que [FileZilla](#), une solution FTP libre. Ce type de solution nécessite de connaître l'architecture du serveur pour retrouver rapidement la donnée recherchée ou de disposer d'une adresse qui nous permet de naviguer facilement dans l'arborescence du serveur de données. Mais la solution visée est de disposer d'un accès aux données parfaitement intégré dans les différents services concomitants, à savoir les catalogues de métadonnées et les outils de consultation/visualisation.

Chaque propriétaire de données décidera librement de leur mise à disposition. Il veillera cependant à respecter les droits afférents à chaque lot d'information en matière de diffusion.

¹⁴ Code de la propriété intellectuelle - art. L123-1 (V)

Un répertoire structuré d'utilisateurs sera mis en place. Pour chaque groupe d'utilisateurs, des droits seront définis quant à l'accès à la métadonnée, à la visualisation de la donnée et à son téléchargement.

Comme cela a été préconisé par Sander Van der Leeuw, Invité Scientifique Annuel du Réseau (ISAR), *la mise en place d'outils communs à tous les OHM pour permettre le partage des données est une priorité* [4^{ème} séminaire annuel du Réseau des OHM, Université d'Aveiro (Portugal), 17-19 Juin 2013]. Cette dernière fonctionnalité devra donc être rapidement développée aux cours de la seconde phase de développement d'OGIIS.

II.2. L'ARCHITECTURE GLOBALE DE LA PLATE-FORME OGIIS

Compte tenu des fonctionnalités attendues et de la nature même du projet, à savoir un outil collaboratif pour des observatoires géographiquement dispersés, la plateforme OGIIS et son architecture devront répondre aux impératifs suivants.

II.2.1. UN HEBERGEMENT UNIQUE, UNE ADMINISTRATION DECENTRALISEE, UN ACCES GENERALISE (EN TOUT LIEU, EN TOUT TEMPS)

Les OHM regroupent de nombreuses équipes scientifiques réparties sur plusieurs continents. Le garant de l'accès à OGIIS sera donc nécessairement internet. L'utilisation simple (consultation et saisie des métadonnées) ou plus complexe d'OGIIS (import/export en lots) devra pouvoir se faire en ligne.

L'administration de la plate-forme (cf. section 1.III.1.) sera partagée entre tous les OHM. Ceci résulte d'échanges menés entre les OHM : les OHM sont aptes à gérer efficacement les données qui sont produites ou acquises dans le cadre de leurs activités de recherche. Toutefois, une supervision de cette administration sera assurée par l'Ingénieur Bases de Données du DRIIHM.

L'hébergement des serveurs supportant les différentes briques logicielles sera, dans la mesure du possible, centralisé en un lieu unique, au plus près de l'Ingénieur Bases de Données rattaché au DRIIHM. Si cela ne s'avère pas réalisable, l'ingénieur veillera à disposer d'un accès facile et régulier à ces espaces et il s'assurera d'avoir une marge d'action suffisante quant au choix du langage et du format des bases de données, au type de sauvegardes à opérer, à la structuration des données et à la définition des accès.

Les différents modules de la plateforme seront accessibles via un guichet unique. Pour le moment, ceux qui existent ou qui sont en phase de développement sont intégrés au site du [ROHM](#).

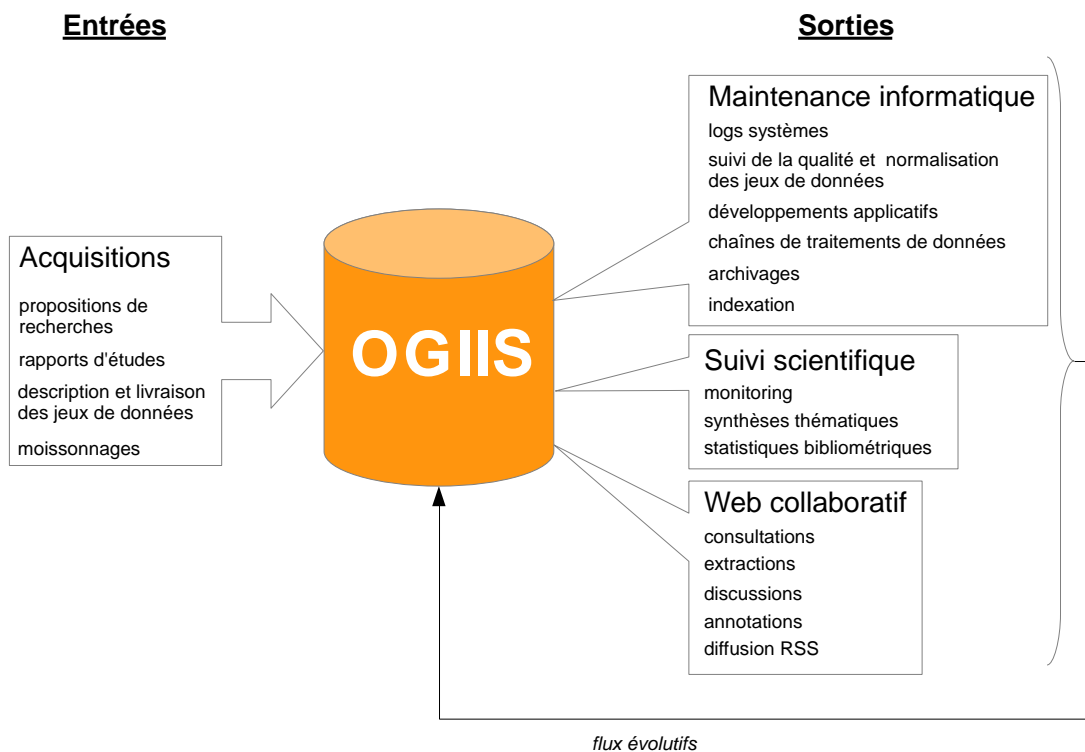


FIGURE 2- LE WEB ET LA PLATEFORME OGIIS

La figure 2 montre la place souhaitée d'OGIIS dans l'écheveau des pratiques numériques. Elle s'inspire des schémas proposés pour la mise en œuvre d'ACMS (Agility Chain Management System) ou chaîne d'agilité [Tassin 2010] dans le traitement de l'information. L'intérêt du concept réside dans la formalisation du circuit dans lequel circulent les données avec des retombées en termes de transparence et de fiabilité en sortie.

II.2.2. UNE STRUCTURE MODULAIRE EVOLUTIVE

OGIIS reposera sur un ensemble de briques répondant chacune à des attentes spécifiques énoncées par le collectif qui s'est constitué autour de ce projet. Ces briques seront développées et assemblées au fur et à mesure de l'avancement du projet. Ce dernier, rappelons-le, dispose du soutien du DRIIHM dont l'activité de recherche évoluera selon la programmation établie pour au moins les 6 prochaines années.

La construction des premières briques sera en accord avec des choix techniques qui ne devront pas empêcher les développements ultérieurs - qu'ils soient ou non traités dans l'étude de faisabilité. Réciproquement, les choix techniques réalisés dans la 2^{ème} phase du projet ne devront pas remettre en cause les décisions opérés lors de la 1^{ère} phase du projet. Le respect de cette stratégie assurera une évolution graduellement progressive à la plateforme OGIIS.

Chaque module pourra reposer sur des solutions techniques ou logicielles différentes les unes des autres si cela signifie des gains de performance. Mais toutes ces briques seront, en revanche, connectées les unes aux autres. Les échanges entre briques seront matérialisés par des flux d'information normalisés. L'indexation des contenus et l'interopérabilité des processus (et notamment leur automatisation par des scripts) serviront une transmission fluide entre les différents modules, comme entre autant de rouages d'une même machine.

Pour préciser davantage quelles seront les modules à construire dans la 1^{ère} phase du projet, voici leur présentation rapide, organisés par ordre de développement dans le temps :

- le référentiel Acteurs/Projets (module 1).
- Le catalogue de métadonnées (module 2).
- Les outils de conversions (module 3).
- Le système d'archivage (module 4).

Les briques à développer au cours de la 2^{ème} phase du projet seront :

- les outils de consultation/visualisation,
- l'accès aux données (avec système de téléchargement).

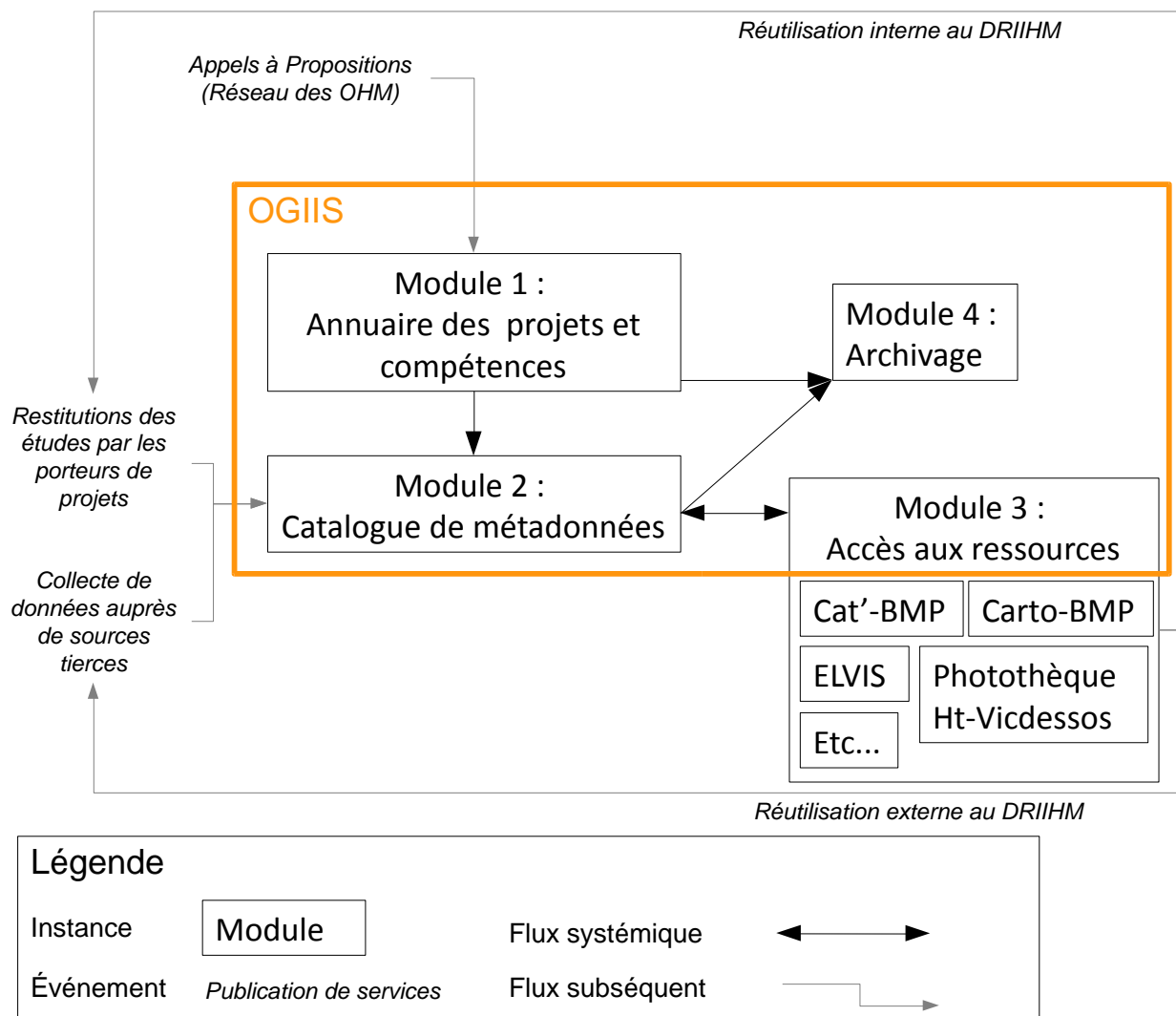


FIGURE 3 - LES MODULES CONSTITUTIFS DE LA PLATEFORME OGIIS

III. LES MODALITES TECHNIQUES ET ORGANISATIONNELLES POUR LA MISE EN PLACE D'OGIIS

La perspective de développement d'une plateforme de gestion de l'information scientifique appelle une réflexion sur les moyens (techniques, humains et financiers) à déployer. S'il est important de s'appuyer sur les forces vives du ROHM afin de se répartir la charge de travail, le recours à des ressources extérieures aux équipes, telles que le LabEx, doit être envisagé comme une opportunité, le cas échéant. Les OHM doivent tous se mobiliser afin de donner une réelle dimension collective au projet et aider à la création d'une identité commune autour d'un outil. Dans ces conditions, au-delà des ressources à allouer au projet, il est nécessaire de construire un système de règles pour assurer la mise en place et la maintenance d'OGIIS. A cette fin, la présente étude de faisabilité vise le dimensionnement des moyens à mettre en œuvre et précise la manière dont ils seront utilisés.

III.1. SPECIFICATIONS ORGANISATIONNELLES

Il est impératif de définir des mécanismes de gestion et des règles pour organiser et répartir les différentes tâches entre OHM et entre les membres d'un même observatoire. Pour y parvenir, dans un premier temps, des groupes d'utilisateurs sont à définir en relation avec les missions qui leur seront confiées dès que l'outil sera opérationnel.

III.1.1. ADMINISTRATEURS ET UTILISATEURS

Dans le cadre du fonctionnement de ce type d'outil, deux grands groupes sont très souvent identifiés :

1) **le groupe des administrateurs** : à ce niveau, il est possible de distinguer les *administrateurs locaux* du *super administrateur* qui sera en charge de la gestion et du développement de la plateforme,

2) **le groupe des utilisateurs** : ils opèreront des recherches de données dans le catalogue (*utilisateur*) et/ou saisiront des fiches de métadonnées pour les lots d'information qu'ils ont/auront produits (*catalogueur*).

LE GROUPE DES ADMINISTRATEURS

▪ Une administration locale assurée par les ingénieurs référents

Les OHM constituent les fondations sur lesquelles s'appuient le ROHM et le DRIIHM. Ils sont à l'interface entre les équipes scientifiques sélectionnées par les APR et les instances de gouvernances supérieures (Conférence des Directeurs, Conseil Scientifique). C'est donc à ce niveau que seront privilégiées l'animation et la gestion quotidienne de la plateforme. Chaque OHM devra définir un ingénieur référent qui supervisera l'utilisation de l'outil par les différents utilisateurs et catalogueurs. Le rôle de ce dernier sera de :

- communiquer sur l'outil auprès des utilisateurs potentiels,
- former les catalogueurs à la saisie,
- former au requêtage tous les usagers potentiels de l'outil,
- valider les saisies effectuées par les catalogueurs
- veiller au bon fonctionnement des passerelles entre OGIIS et les outils déjà existants dans son OHM. En cas de dysfonctionnement, il devra s'engager à faire remonter, dans les meilleurs délais, le problème au super administrateur,
- faire part des nouveaux besoins en termes de développements de la plateforme auprès du super administrateur.

▪ Une super administration au niveau du DRIIHM

L'administration globale de la plateforme relève directement des attributions du DRIIHM. C'est à ce niveau que sera basé le super administrateur de la plate-forme qui aura pour responsabilité d'assurer le bon fonctionnement d'OGIIS, de réfléchir avec les équipes à son évolution, d'aider au développement de nouvelles fonctionnalités et de soutenir les observatoires dans leurs missions en lien avec OGIIS. Plus concrètement, ses principales missions seront de :

- réaliser la 1^{ère} phase de développement d'OGIIS aux regards des éléments exposés dans la présente étude de faisabilité. Pour cette tâche, il pourra en partie s'appuyer sur les compétences présentes au sein du réseau et sur les partenaires proches (ex : l'UMS BBEES).

- Assurer l'interfaçage entre les instances de gouvernance du DRIIHM et les ingénieurs référents. Ainsi, il procédera à la collecte des besoins en développement auprès des ingénieurs référents et il les fera ensuite valider par les instances de décision du DRIIHM.
- Mettre en œuvre les développements ultérieurs et en premier lieu, les éléments de la phase 2, incluant le module de « consultation/visualisation » et le module de « stockage/téléchargement ».
- Former et assister, si besoin, les ingénieurs référents à l'utilisation de la plateforme.

LE GROUPE DES UTILISATEURS

▪ Des catalogueurs et des utilisateurs

Partant du principe que la personne la plus à même de décrire une ressource est son producteur, la saisie des fiches de métadonnées sera à la charge des responsables¹⁵ des études financées au titre des APR. Ces derniers devront s'assurer que chaque lot de données produit au cours d'une étude financée par le DRIIHM ait une fiche de métadonnée associée. Les scientifiques en charge de la saisie constituent le groupe des catalogueurs et ont, à ce titre, un accès privilégié à la plateforme par rapport aux utilisateurs simples. Ces derniers constituent un groupe « grand public » : étudiants, chercheurs hors OHM, gestionnaires et opérateurs, bureaux d'études... La plateforme se doit d'offrir un accès large pour, au moins, la consultation des métadonnées.

III.1.2. INTERACTIONS ENTRE LES DIFFERENTS GROUPES ET ENTRE LES INSTANCES DE DECISION DU DRIIHM ET DU ROHM

La gestion et l'administration d'OGIS impliquera de nombreuses interactions entre les différents groupes. *Il est impossible* d'identifier, dès à présent, l'ensemble de ces interactions. Mais, pour le bon fonctionnement de la plateforme, quelques repères organisationnels doivent être énoncés. Le schéma ci-dessous les illustre.

¹⁵ Pour chaque étude financée par les APR des OHM ou du ROHM, un ou plusieurs responsable d'étude sont spécifiés, cf. item 2 « Responsable » de la demande de financement des APR.

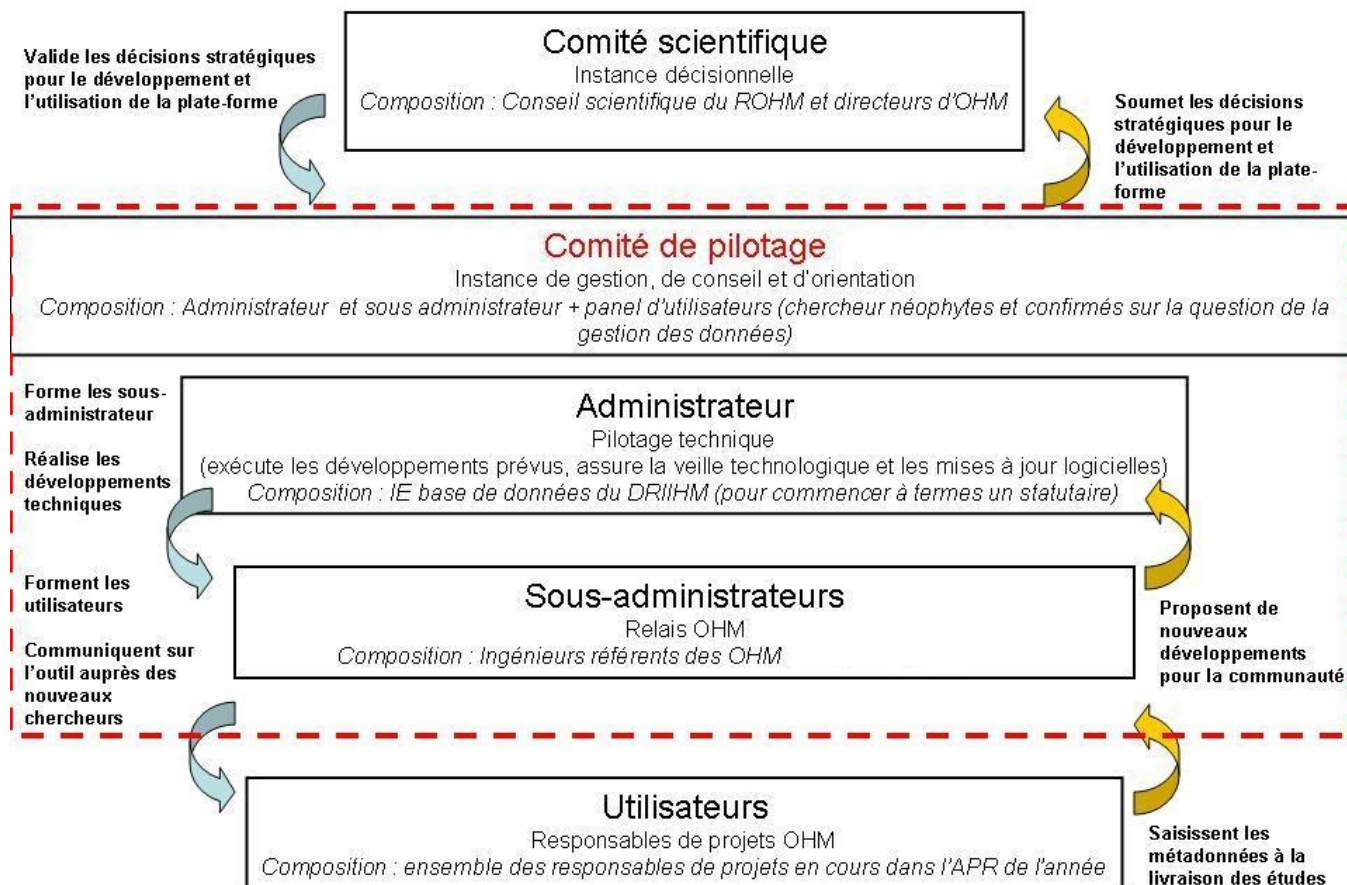


FIGURE 4 - INTERACTIONS ENTRE LES DIFFERENTS PROTAGONISTES D'OGIIS

Les points suivants mettent l'accent sur les interactions à ne pas négliger.

- **Interactions entre les administrateurs et les utilisateurs**

Chaque OHM sera autonome quant à l'organisation de ses sessions de formations auprès des utilisateurs. Il pourra cependant s'appuyer sur les compétences du super administrateur rattaché au DRIHM, sur les compétences de leur ingénieur référent et sur celles des deux correspondants INSPIRE que compte le ROHM : Amandine BLIN (UPR 2147 DEH) et Maxime MIGNAN (UMR 7263 IMBE).

Les administrateurs locaux (ou ingénieurs référents) devront, par ailleurs, s'assurer de la saisie annuelle, par leurs catalogueurs, des données produites au sein des projets retenus et financés au titre des APR. Dans la pratique, les résultats des études sont restitués au bout d'1,5 an après le lancement du projet - même si, occasionnellement, des délais supérieurs sont accordés. La mise en forme des données issues d'études scientifiques peut être longue. Un délai maximum de 2 ans après le lancement de l'étude est donc jugé raisonnable pour permettre la saisie des métadonnées. Au-delà de ce délai, l'OHM doit être en mesure d'exercer un droit de regard auprès des porteurs des projets qu'il a financé de telle sorte que les délais de restitution soient, le cas échéant, justifiés et adaptés. En accord avec la programmation scientifique du DRIHM et l'engagement que constitue la signature de la charte des OHM, la rétention manifeste de données et à plus

forte raison le non-renseignement des métadonnées peuvent constituer un motif de non-financement si les directeurs d'OHM concernés le juge approprié.

▪ **Interactions entre les administrateurs : les réunions techniques**

Les échanges entre les ingénieurs référents et le super administrateur relèveront de la gestion courante de la plateforme. Il semble important de prévoir, une à deux fois par an, une réunion rassemblant l'ensemble des administrateurs. Cette réunion sera un lieu d'échanges techniques à propos de la plateforme et de son utilisation au quotidien. Elle constituera un temps pour la remontée des besoins des équipes en développements.

Du personnel technique extérieur au ROHM pourra être convié à ces réunions : il viendra enrichir les réflexions en faisant part de sa propre expérience.

Ces interactions sont nécessaires ; elles permettront d'entretenir une dynamique collective autour du projet OGIIIS.

▪ **Interactions entre les différents groupes d'utilisateurs identifiés : le comité de pilotage**

L'idée de la création d'un comité de pilotage est apparue au cours des différentes réunions organisées dans le cadre du projet OGIIIS. Ce comité aurait pour objectif, comme pour les réunions techniques entre administrateurs, d'impulser une dimension collective au projet OGIIIS en l'élargissant aux utilisateurs et aux directeurs. Le besoin de réunir tous les intervenants de la plateforme est évident, au moins pour collecter les retours d'expérience de l'ensemble des utilisateurs de la plateforme.

Les porteurs de projet souhaitent que ce comité de pilotage réunisse différents publics : **des experts** habitués à la manipulation des données et des métadonnées et **des néophytes**, c'est-à-dire des personnes qui sont motivées par cette question de la gestion de la donnée à caractère scientifique, mais qui n'ont pas de connaissance ou de compétences *a priori* sur le sujet. Ce comité ne doit pas être trop important en termes de participants (ce qui complexifierait son organisation). L'idée de constituer le comité de pilotage, en lien avec le séminaire annuel du ROHM, a émergé à Aveiro (juin 2013) : elle doit encore être validée et mise en pratique.

III.2. LES SPECIFICATIONS TECHNIQUES

III.2.1. TYPE DE DONNEES ET CHOIX DES NORMES DE DESCRIPTION

Les OHM couvrent des champs disciplinaires larges (cf. section 1.1.1). Il en résulte des types de données et des formats de données variés. Suite à la réunion OGIIIS de Lyon, le 17 Avril 2013, le collectif a souhaité rester sur des choses simples plutôt que de réfléchir à une ontologie¹⁶ des catégories de données au niveau du ROHM. Il a décidé également de partir sur une liste large des catégories de données potentiellement gérées par les OHM. Dans un 2^{ème} temps, après une phase de saisies de métadonnées sur un temps qui reste à définir, les appellations et le contenu des catégories seront reprécisés si nécessaire. Enfin, parce que tous les OHM ne sont

¹⁶ **Une ontologie** est un ensemble structuré des termes et de concepts représentant le sens d'un champ d'informations, que ce soit par les métadonnées d'un espace de noms, ou les éléments d'un domaine de connaissances. L'ontologie constitue en soi un modèle de données représentatif d'un ensemble de concepts dans un domaine, ainsi que des relations entre ces concepts. Elle est employée pour raisonner à propos des objets du domaine concerné.

pas forcément habitués à manipuler certains types de données ou parce que les vocables n'ont pas forcément la même définition pour tous les OHM, à chaque catégorie de données retenue, des descripteurs ont été associés afin de mieux expliciter ce que regroupe chacun de ces termes. L'harmonisation sémantique pourra être poursuivie par la personne en charge de la maintenance de l'outil.

- ❖ **Données Raster** : donnée image organisée sous la forme d'une grille régulière, dont l'élément structurant est le pixel. Chaque pixel (ou cellule de grille) couvre une aire géographique donnée et une valeur d'attribut est assigné à la cellule. Des exemples de données raster : une image satellite, le SCAN25...
- ❖ **Données vectorielles** : les éléments géographiques représentés par des données vectorielles sont localisés par des coordonnées (X, Y, (Z)). Ils sont constitués de trois types d'entités : point, ligne, polygone.
- ❖ **Cartes et atlas** : un **atlas** est un recueil de cartes géographiques, conçu pour représenter un espace donné et exposer un ou plusieurs thèmes (géographie, économie, histoire...). Une **carte** fournit une représentation graphique du monde réel qui permet de voir la localisation d'objets ou de phénomènes. Elle peut être géoréférencée et répondre à une mise en page normalisée (échelle, légende, source...).
- ❖ **Ressources documentaires** : elles incluent la **bibliographie grise** (communication orale, rapport de stage, rapport d'expertise, presse...) et celle qui est **référéncée** (articles, ouvrages, etc.).
- ❖ **Corpus de textes** : ensemble d'énoncés retenus, écrits ou oraux, qui sera soumis à l'analyse. Il peut inclure les éléments suivants : articles de presse, brochures de communication...
- ❖ **Enquêtes/Entretiens** : l'entretien est une technique de recueil qui engage deux personnes en vis-à-vis et à ce titre ne peut être considéré comme un simple questionnaire où on est dans une relation anonyme. **NB** : cette catégorie de données a été individualisée de la catégorie précédente « corpus de textes » parce qu'un entretien contient souvent des données personnelles permettant d'identifier la personne rencontrée et/ou des données sensibles.
- ❖ **Données tabulaires** : données organisées sous la forme d'un tableau ou selon une syntaxe informatique structurante (lignes, colonnes, séparateurs de champs) : résultats d'analyses, chroniques de données, listes d'enregistrements, inventaires...
- ❖ **Programmes/Codes/Scripts** : un **programme** est une séquence d'instructions qui spécifie, étape par étape, les opérations à effectuer pour obtenir un résultat. Un **code** est en ensemble de règles définissant une correspondance univoque entre des informations et leur représentation par des caractères, des symboles ou des éléments de signal. Un **script** est un programme informatique qui ne nécessite pas de compilation avant d'être exécuté. Pour fonctionner, les scripts doivent être interprétés par un programme ou un serveur dédié au langage dans lequel ils ont été écrits. Tous (Programmes/Codes/Scripts) sont destinés au traitement automatique de la donnée.
- ❖ **Audio/Vidéo** : fichiers audiovisuels ou vidéos.
- ❖ **Photographies/Iconographies** : les **photographies** sont des images obliques, verticales ou horizontales n'ayant pas été redressées des erreurs de topographie et n'ayant pas été géoréférencées. Les **iconographies** sont des représentations artistiques d'un sujet sur différents supports (ex : gravures, peintures, sculptures, lithographies, ...).
- ❖ **Schémas/Graphiques/Figures** : un **schéma** est une représentation graphique réduite à l'essentiel, et souvent symbolique, mais où toutes les informations se trouvent données de façon précise. C'est une représentation des constituants fondamentaux d'un objet complexe, incluant les relations fonctionnelles existant entre ces constituants ou bien une représentation de la structure fondamentale d'un phénomène, d'une organisation. Un **graphique** est une représentation symbolique ou à coordonnées cartésiennes d'un phénomène statique ou en mouvement observé en fonction d'un

autre paramètre (durée, coût...). Une **figure** est une représentation dans le plan ou dans l'espace d'une forme correspondant à une abstraction ou à une action organisée.

- ❖ **Autre(s)**: cette catégorie regroupe toutes les catégories qui auraient pu être omises lors des échanges entre OHM. A terme, si nombre de données viennent à être classées dans cette catégorie, il sera important de réfléchir à la création de nouvelles catégories.

Les trois premiers types de données appartiennent à la catégorie des **données géographiques**, définies comme étant « *des données faisant directement ou indirectement référence à un lieu ou à une zone géographique spécifique* » (*Directive Inspire, article 3*). Ce type de données a été largement considéré lors des échanges menés sur l'amélioration des pratiques puisqu'il est directement concerné par la **Directive INSPIRE** (Infrastructure for Spatial Information in Europe). Cette dernière, entrée en vigueur le 15 Mai 2007 et transposée en droit français en 2011, vise à fixer *les règles générales destinées à établir l'infrastructure d'information géographique aux fins des politiques environnementales communautaires, au sein de la Communauté Européenne*. La Directive INSPIRE est un texte légal et général qui fixe des objectifs. Elle est complétée par des **règles de mise en œuvre**. Ces dernières imposent des obligations et indiquent des moyens. Une des règles de mise en œuvre concerne les métadonnées de données : 19 champs obligatoires sont spécifiés. Les métadonnées INSPIRE sont basées sur la norme ISO 19115 (jeux de données) et ISO 19119 (services), mais elles intègrent également des éléments étendus et différents des éléments obligatoires de la norme ISO établie par l'Organisation Internationale de Normalisation.

Certains OHM ont d'ores et déjà choisi d'exploiter cette directive pour décrire leurs lots de données. Cependant, son usage exclusif n'est pas forcément adapté à l'ensemble des OHM pour les raisons données ci-après :

- le nombre élevé de champs à remplir et la complexité des informations rendent l'usage de cette norme difficile pour les non-initiés,
- cette norme est conçue pour la description de données géographiques. Si potentiellement une information numérique peut être localisée, relève-t-elle alors nécessairement de la directive INSPIRE ? Cet exercice fait encore débat.
- D'autres normes spécialisées sont parfois plus adaptées comme par exemple l'EML ou le Darwin Core¹⁷ (données écologiques), l'EXIF ou l'IPTC¹⁸ (données photographiques), etc.

La présente étude préconise l'utilisation à *minima* d'une norme générique : le **Dublin Core**¹⁹ (DC). Cette dernière est beaucoup utilisée pour la description de ressources telles que les pages web. Elle comprend une liste de 15 champs optionnels et répétables. Ce qui lui confère une grande souplesse d'utilisation. La liste des champs de la norme DC en précisant le texte attendu est en annexe du présent rapport.

¹⁷ EML (Ecological Metadata Language) : il s'agit d'une norme de données élaborée par l'écologie pour l'écologie. Elle est issue d'une initiative du NCEAS (National Center for Ecological Analysis and Synthesis). Elle est d'ores et déjà bien utilisée aux Etats-Unis. Elle peine davantage à s'introduire en France. EML apparaît notamment comme composante du Darwin Core (DwC) dans les préconisations du Gbif [GBIF 2012].

¹⁸ Il y a deux standards principaux pour la description des métadonnées des photographies réalisées à l'aide d'appareil numériques : l'EXIF (http://fr.wikipedia.org/wiki/M%C3%A9tadonn%C3%A9es_EXIF) établi par la Japan Electronic Industry et l'IPTC (<http://www.iptc.org/site/Home/>) édité par l'International Press Telecommunication Council.

¹⁹ Un certain nombre d'équipes impliquées dans la sémantique sur le Web ont organisé un atelier à Dublin (Ohio) en 1995. Ils ont défini une liste précise de quelques métadonnées liées aux pages Web. Ce jeu de métadonnées a été surnommé « Dublin Core metadata ». Désormais, l'ensemble de ces « propriétés » cataloguées par la **DUBLIN CORE METADATA INITIATIVE** (DCMI) comprend plusieurs dizaines de termes, mais 15 d'entre elles, plus fondamentales, ont été normalisées (ISO 15836-2003, RFC 5013). Pour plus d'information voir <http://dublincore.org/>.

L'obligation faite par le cadre réglementaire aux établissements publics est de recenser les données géographiques à caractère environnemental en leur possession. Compte tenu des métadonnées déjà renseignées, en conformité avec cet impératif, la norme INSPIRE s'impose comme un standard de description au niveau du ROHM. Elle est incontournable pour les données géographiques et elle sera utilisée conjointement à la norme DC.

Ces deux formats de métadonnées pourront cohabiter. Des outils pour passer d'un modèle de description à l'autre seront proposés au cours de la phase 1 du développement OGIIS. En attendant les premiers développements, ***un outil d'aide à la saisie des métadonnées, produit par le BRGM et adapté par les ingénieurs participant à OGIIS, est à présent accessible aux OHM. Il permet la description des lots d'information sous les deux formats : INSPIRE et DC.*** Pour information, les principales correspondances sémantiques entre ces deux normes ont été étudiées comme le montre le tableau annexé au présent rapport.

D'autres normes de descriptions de métadonnées pourront être progressivement intégrées dans OGIIS – si des demandes motivées allant dans ce sens sont émises par certains OHM. Pour le moment, très peu de ressources sont décrites selon d'autres normes. Des passerelles de conversion seront cependant envisagées en phase 2 du projet. Les autres normes semblant présenter un intérêt sont prioritairement celles relatives aux ressources photographiques (EXIF et IPTC) et aux données de type inventaires biologiques (Darwin Core²⁰). Le standard international Observation and Measurement (O&M),²¹ issu des travaux de l'Open Geospatial Consortium, est un standard international qui définit un schéma d'encodage pour les observations scientifiques et les objets qu'elles génèrent. Dérivé de trames génériques et basé sur du XML, il s'applique particulièrement (mais pas exclusivement) aux données SIG. Il est lié aux activités d'acteurs internationaux de la standardisation tels que W3C, OGC, GEOSS. Cette large audience et les retombées technologiques potentielles concernant un sujet aussi important pour les OHM plaident en faveur d'une meilleure connaissance de ces standards afin de consolider nos choix conceptuels et sémantiques. A noter que le recensement de ces normes, bien qu'avancé, est toujours en cours et une veille sur cette question devra être assurée par le super administrateur.

III.2.2. CHOIX DE LA (OU DES) SOLUTION(S) TECHNIQUE(S)

Cette question a été notamment abordée au cours de la réunion élargie durant laquelle différents ingénieurs, hors OHM et ayant travaillé sur la mise en place de catalogues de métadonnées, sont venus présenter leurs réalisations. Deux OHM disposent également d'une expérience en la matière (3 instances installées : deux pour l'OHM VR et une pour l'OHM BMP). Enfin, nous nous sommes également intéressés aux choix établis par d'autres réseaux comme celui de l'inter-ZA²² ou d'ANAEE (ANALysis and Experimentation on Ecosystems).

L'avantage principal d'une solution « clef en main » est son coût réduit en termes de développement ; il se limite au temps nécessaire à l'installation de l'application sur un serveur. Exploiter un outil livré prêt à l'emploi offre la garantie d'une interopérabilité maximum entre les différentes installations faites de ce même produit. Cependant, certains inconvénients peuvent aussi voir le jour :

²⁰ Voir <http://rs.tdwg.org/dwc/> et <https://code.google.com/p/darwincore/>

²¹ http://en.wikipedia.org/wiki/Observations_and_Measurements

²² Réseau des Zones Atelier, grands équipement de l'INEE au même titre que les OHM. Pour plus d'information : <http://www.za-inee.org/-Presentation->

- **l'ergonomie** : un logiciel du marché (qu'il soit sous licence free, Open Source ou payante) ne dispose pas forcément d'une interface facile d'accès pour les utilisateurs non-initiés ;
- **l'installation et les mises à jour** : selon le type d'installation (locale ou réseau), selon l'OS du serveur et selon bien évidemment la qualité de développement du produit, l'installation s'avère parfois compliquée. Disposer d'un personnel qualifié devient alors nécessaire à la fois dans l'étape de déploiement de l'application et dans celle de la maintenance à assurer régulièrement pour permettre le bon fonctionnement du logiciel ;
- **les fonctionnalités** : si certaines viennent à manquer dans une application livrée clé en main, il devient alors impossible ou difficile de réaliser, en interne, le développement du module souhaité. En effet, celui-ci nécessite du personnel qualifié et/ou un accès au code source.

Concernant les systèmes reconfigurables ou développés en interne, le TGE Adonis a montré l'intérêt de ce type de solutions dont la grande plasticité (évolutivité et ergonomie) est basée sur le parti pris de s'affranchir des normes pour imposer un minimum de contraintes à l'utilisateur et ainsi atteindre une audience et une emprise plus larges

La ZABR, Zone Atelier pionnière sur les questions de géocatalogue, a développé sa propre solution logicielle complètement adaptée aux données dont elle avait la charge dès 2002. La conception et le développement de l'outil ont nécessité beaucoup de temps. Du fait de son développement « maison », malgré tous les soins qui y ont été apportés, procéder à une évolution du produit nécessitait du temps de chercheurs important. Une des critiques majeures qui a été faite à ce produit est, en plus de son évolution limitée, sa cartographie pseudo-dynamique. Depuis lors, la ZABR a choisi de remplacer son outil « maison » par un outil disposant d'une communauté de développeurs active, à savoir GeoNetwork. Mais la difficulté de ce passage d'une solution développée en interne à une solution plus aboutie est de s'adapter aux nouvelles fonctionnalités, à la logique du nouvel outil et surtout, d'assurer le transfert des données sans perte. L'adaptation n'est pas toujours aisée.

D'autres initiatives, plus lointaines mais non moins intéressantes, montrent la faisabilité d'un catalogage normalisé basé sur des systèmes de publication de contenu génériques²³ (à savoir le CMS pour Content Management System). Cette approche modulaire (orientée plugins) privilégie un rehaussement de l'existant (tous les OHM ont au moins publié un CMS ou prévoient de le faire) en termes de fonctionnalités sur l'installation de nouvelles applications. Autrement dit, compte-tenu de la diversité des tâches requises pour la coordination d'un OHM, cette solution se présente comme un scénario d'évolution continue, répondant aux besoins, au fur et à mesure qu'ils se présentent. Rattachés à une base commune, ces développements, implémentables séparément les uns des autres, seraient plus facilement mutualisés que le recours ad hoc à des solutions spécialisées supposées plus coûteuses en temps et en infrastructure informatique.

La présente étude met en avant les critères d'évolution (sur le long terme) et de portabilité dans le choix de solutions logicielles. Ces dernières devront permettre les développements spécifiques énoncés par le ROHM et le LabEx DRIIHM. **A l'heure actuelle, le collectif s'orienterait davantage vers le développement de sa propre solution de gestion de métadonnées en accord avec la norme Dublin Core.** En effet, les systèmes existants (tels que GeoNetwork) sont jugés encore trop complexes pour attirer les chercheurs vers ce type d'activité. En revanche, pour ce qui est de la gestion de contenu, tels que les photothèques (Piwigo) ou les outils de gestion de ressources documentaires (Zotero) pourraient faire appel à des outils, Open Source le plus souvent, comme ceux cités entre parenthèses. Ces outils sont d'ores et déjà exploités par un ou des OHM ou directement par le

²³ <http://lab.usgin.org/groups/drupal-development/creating-iso-19139-metadata-through-drupal-views-and-views-bonus-pack>

CNRS. Concernant la gestion des données à référence spatiale, certains OHM disposent d'un vrai savoir-faire qui permettrait de favoriser une solution Open Source qui est GeoNetwork. Mais pour l'heure, il est encore difficile de faire des choix définitifs quant aux solutions à retenir. Chacun soutient sa solution sans avoir une expertise avancée dans l'évaluation des autres solutions du marché. Dans ces conditions, **au moins une personne du ROHM sera identifiée par outil de gestion souhaitée par le collectif. Celle-ci sera chargée de définir les besoins des équipes en la matière et d'évaluer les solutions existantes pour savoir si l'une d'elles peut répondre aux attentes du groupe.** Dans le cas contraire, une solution sera développée en interne. Ce travail d'inventaire, de collecte des besoins et éventuellement de conceptualisation se fera en lien avec l'Ingénieur Bases de Données du DRIIHM. Voici donc les outils qui sont envisagés au sein d'OGIIS au moment où se finalise cette étude, ainsi que le nom de la (ou des) personne(s) qui seront en charge de l'étude.

Type de services	Personnels	Propositions techniques
Sites Web	Maxime Maignan A contacter : Armand Rotereau	CMS (SPIP)
Photothèque	Emilie Lerigoleur Franck Vidal	Piwigo (+ CMS)
Cartographie dynamique	Guillaume Fantino Émilie Lerigoleur Maxime Maignan Kristell Michel Jean-Claude Raynal A contacter : Armand Rotereau	GeOrchestra, PostgreSQL, PostGIS, Openlayers
Géocatalogue	Guillaume Fantino Émilie Lerigoleur Maxime Maignan Kristell Michel Jean-Claude Raynal A contacter : Armand Rotereau	Geonetwork, Geosource, GeOrchestra, CMS
Ressources Documentaires	Maxime Maignan A contacter : Damien Davy Armand Rotereau	Zotero (+ CMS), DSpace, Koha, VuFind

Tableau 1 : Propositions de répartition des ingénieurs référents par services

Pour les OHM qui disposent déjà de solutions pour gérer certains lots de données, avec la création d'OGIIS, ils pourront choisir d'exploiter les nouveaux outils mis à leur disposition et/ou conserver leur(s) outil(s) et assurer le transfert de leurs métadonnées par moissonnage. Ainsi, une seule saisie de l'information sera utile quelle que soit la solution retenue.

III.2.3. AUTHENTIFICATION ET ACCES A LA PLATEFORME OGIIS

Le service de consultation ouvert au grand public sera limité aux métadonnées, aux données libres de droits ou sous licence ouverte et à certaines informations provenant du référentiel « acteurs et projets » permettant de savoir qui sont les membres des OHM et ce qu'ils font dans ce cadre.

Les autres fonctionnalités offertes par OGIIIS seront soumises à l'authentification de l'utilisateur. Cette dernière sera possible pour toutes personnes investies dans un ou plusieurs OHM, ainsi que les personnes travaillant pour le collectif (Conseil Scientifique, DRIIHM, ROHM). Ce système d'authentification sera unique pour l'ensemble des services proposés par les OHM. Il sera centralisé et administré par le super administrateur. Il permettra, par exemple, de s'identifier sur le site du ROHM et éventuellement sur ceux des OHM - qui restent toutefois libre de disposer d'un système d'identification propre. Le système d'authentification devra être mis à jour sur demandes des responsables d'OHM et au moins deux fois par an, suite aux sélections des projets et des équipes de recherche correspondantes.

Les utilisateurs identifiés comme responsable d'étude auront accès à la gestion des informations relatives à leur projet dans le module 1 « référentiel Acteurs/Projets » (dépôt/modification de projet pour l'APR, cf. sous-section III.3.1, mais également au module 2 « catalogue des métadonnées » pour saisir la description, en fin de projet, des lots de données produits au cours de l'étude.

Les administrateurs (ou ingénieurs référents) et le super administrateur auront les privilèges leur permettant de réaliser les tâches d'administration qui leur reviendront.

III.2.4. MATERIELS REQUIS ET HEBERGEMENT

Une infrastructure de recherche comme celle-ci nécessite des moyens techniques à déployer rapidement pour lancer la 1^{ère} phase de développement du projet. Ces moyens seront susceptibles d'évoluer au fil du temps (mises à jour, remplacement des serveurs, développements prévus en phase 2 et ceux qui suivront). Il est donc nécessaire de prévoir des budgets de fonctionnement récurrents pour la maintenance et l'amélioration de la plateforme sur une période minimum de 6 ans (durée du DRIIHM). A titre d'information, 17% des budgets annuels du Long Term Ecological Research Network²⁴ sont consacrés à l'informatique et à la gestion des bases de données.

La construction de la plateforme et son accessibilité via le web demande la mise à disposition de serveurs. Pour les besoins de développement de la phase 1, une distinction doit être faite entre la base de données référentiel Acteurs/Projets, le catalogue de métadonnées et l'annuaire d'identification LDAP²⁵ (Lightweight Directory Access Protocol). Cette distinction concerne la conception et l'hébergement : ces outils bien qu'étant appelés à interagir différemment par leurs usages et leurs caractéristiques techniques. **Pour chaque outil, on doit disposer d'une instance de développement et d'une instance de production de manière à assurer la continuité de services durant les opérations de maintenance.** La continuité des services en ligne serait compromise si cet élément venait à être négligé.

La question de l'hébergement des serveurs a été discutée, notamment avec l'UMS 3648 BBEES. Cette unité de service prévoit une demande globale d'hébergement pour les besoins de l'INEE auprès du Centre de Calcul de l'IN2P3 (Institut National de Physique Nucléaire et de Physique des Particules) : il s'agit d'une autre unité de service et de recherche du CNRS, classée parmi les grandes infrastructures françaises de recherche, qui se doit d'offrir notamment des services en matière de stockage et de traitement de grandes masses de données, ainsi que leur transfert sur des réseaux internationaux à très haut débit. En attendant la mise à disposition de cet hébergement, prévue pour 2016, les services proposés par la DSI du CNRS ou celles

²⁴US Long Term Ecological Research Network développe depuis sa création en 1980 une approche de l'étude des changements socio-écologiques basée sur le choix de sites spécifiques et sur la collecte de données sur de longues périodes. Pour en savoir plus, voir <http://www.lternet.edu/> et [LAGADEUC & CHENORKIAN 2009]

²⁵ Le LDAP est, à l'origine, un protocole permettant l'interrogation et la modification des services d'annuaires. Voir http://fr.wikipedia.org/wiki/Lightweight_Directory_Access_Protocol et Section 3.03(a) du présent document.

proposées par la TGIR Huma-Num sont susceptibles de répondre aux besoins pour le développement de la phase 1.

III.3. PROPOSITION DE PLANNING ET RECOMMANDATIONS POUR LA POURSUITE DU PROJET

Deux phases principales sont donc prévues pour le développement de la plateforme OGIIS :

- **PHASE 1** : développement et mise en production des modules 1. Référentiel Acteurs/Projets, 2. Catalogue de métadonnées, 3. Outils de conversions, 4. Système d'archivage
- **PHASE 2** : développement des modules Consultation/Visualisation et Stockage/Téléchargement.

La phase 1 (cf. tableau 2) a déjà débuté dans la mesure où le module **REFERENTIEL ACTEURS/PROJETS** est déjà en cours de construction grâce à la collaboration de l'OHM HV et du ROHM. Cependant, les réalisations prévues dans cette phase sont fonction de deux facteurs limitant :

- les moyens techniques nécessaires aux développements informatiques,
- le(s) recrutement(s), notamment celui de l'ingénieur d'études, chargé des bases de données du DRIIHM²⁶ et dont la participation aux développements et à l'administration d'OGIIS est fondamentale.

Le planning exposé ci-dessous dépend de ces deux points :

Légende : **En rouge**, les éléments se rapportant à la phase 1. **En bleu**, ceux qui seront développés en phase 2 du projet.

2013	2014		
4 ^{EME} TRIMESTRE	1ER TRIMESTRE	2EME TRIMESTRE	3EME TRIMESTRE
- développement du référentiel Acteurs/Projets	- recrutement d'un IE BDD DRIIHM - mise en production du référentiel Acteurs/Projets pour l'APR 2014	- développement du Catalogue de métadonnées - communication autour du référentiel Acteurs/Projets au cours du séminaire ROHM 2014 - composition du comité de pilotage au séminaire ROHM	- finalisation du Catalogue de métadonnées - développement des passerelles outils OHM/OGIIS - réflexion préliminaire pour la mise en place du module Archivage

²⁶ Le profil technique de cette personne appelle des compétences en informatique et gestion de bases de données (+ des compétences en web pour le développement des interfaces hommes/machine). Profil type CNRS : E2A21.

2014	2015		
4 ^{EME} TRIMESTRE	1ER TRIMESTRE	2EME TRIMESTRE	3EME TRIMESTRE
<ul style="list-style-type: none"> - test du module Catalogue de métadonnées - formation des ingénieurs référents à OGIIIS - développement du module Archivage 	<ul style="list-style-type: none"> - implémentation de l'existant dans le module métadonnées - utilisation du référentiel Acteurs/Projets pour l'APR 2015 	<ul style="list-style-type: none"> - test et mise en service du module Archivage - formation des catalogueurs - communication lors du séminaire ROHM 2015 	<ul style="list-style-type: none"> - réflexion préliminaire sur le lancement de la phase 2

2015	2016		
4EME TRIMESTRE	1ER TRIMESTRE	2EME TRIMESTRE	3EME TRIMESTRE
<ul style="list-style-type: none"> - développement du module de Consultation/ Visualisation 	<ul style="list-style-type: none"> - saisies des métadonnées de l'APR 2015 - mise en service du module Stockage/ Téléchargement - réflexion préliminaire sur le module Stockage/ Téléchargement 	<ul style="list-style-type: none"> - développement du module Stockage/ Téléchargement 	<ul style="list-style-type: none"> - mise en service du module Stockage/ Téléchargement

Tableau 2 : Planning prévisionnel pour la mise en place d'OGIIS

PARTIE 2. METHODES, RESULTATS ET PRECONISATIONS POUR LA MISE EN PLACE D’OGIIS

Cette partie est une présentation des résultats produits par la démarche entreprise dans le projet OGIIS. Ces résultats sont des éléments techniques et constituent les premières pierres avec lesquelles les développements peuvent démarrer.

I. METHODOLOGIE DE TRAVAIL POUR LE PROJET OGIIS

I.1. DEROULEMENT DU PROJET

L'étude de faisabilité comporte deux temps forts. Tout d'abord, une phase d'échanges et de concertation s'est engagée, entre les différents OHM, afin de définir les attentes et les contraintes de chacun en termes de gestion de la donnée. Le deuxième temps du projet est dédié à des échanges techniques entre demandeurs et concepteurs.

I.1.1. PHASE 1

La phase 1 a consisté globalement à définir les besoins des futurs utilisateurs au sein des OHM et à lister les solutions techniques existantes - éventuellement déjà testées par certaines équipes. Une phase 1.1 a permis de lister les outils existants dans les différents OHM, tandis qu'une phase 1.2 a précisé les attentes complémentaires des OHMs par rapport à la future plateforme, en lien avec les retours d'expérience sur des outils déjà existants dans d'autres structures de recherche. La définition de ces attentes a porté notamment sur les types de données/métadonnées à gérer, les niveaux de sécurité et les fonctionnalités de la plateforme à prévoir. Nous avons aussi répertorié les points critiques ou points de blocage possibles pouvant gêner la bonne poursuite du projet. A l'issue de cette phase un document de travail a été produit : un cahier des charges fonctionnel, décrivant les fonctions attendues de la plateforme et précisant les spécifications techniques des besoins exprimés.

I.1.2. PHASE 2

Au cours de la 2^e phase du projet, nous avons précisé les spécifications techniques (choix des logiciels, des formats, des normes à respecter) pour la future plateforme. Des échanges, des négociations et des compromis ont été arbitrés entre les demandeurs (les OHM) et les « concepteurs » (en l'occurrence l'interlocuteur principal a été l'UMS BBEES qui supervisera le travail de l'ingénieur Bases de données du Labex DRIIHM). Cette phase 2 a pris la forme de réunions régulières entre les différents partenaires du projet. A l'issue de cette phase, un document, le présent dossier de définition ou étude de faisabilité, a été produit. Ce dernier contient l'ensemble des préconisations techniques et organisationnelles (modalités de coopération, budgétisation, moyens humains) à prendre en compte pour la mise en place de la plateforme. Enfin, la phase finale du projet sera une phase de restitution du projet (qui a déjà en partie eu lieu au séminaire du ROHM à Aveiro) et la

présentation de la plate-forme aux futurs utilisateurs (à prévoir au cours dans la phase 2 de développement, cf. section III.3.).

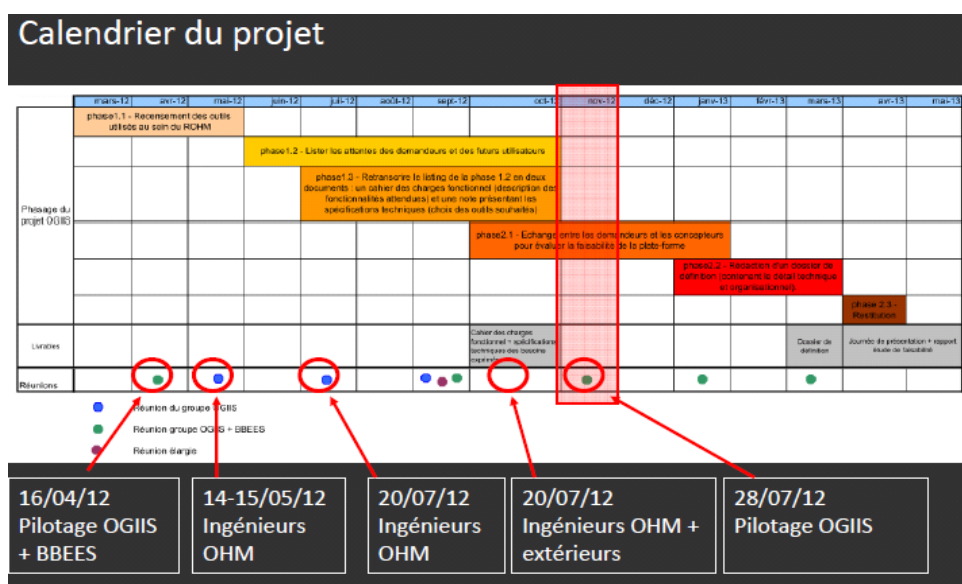
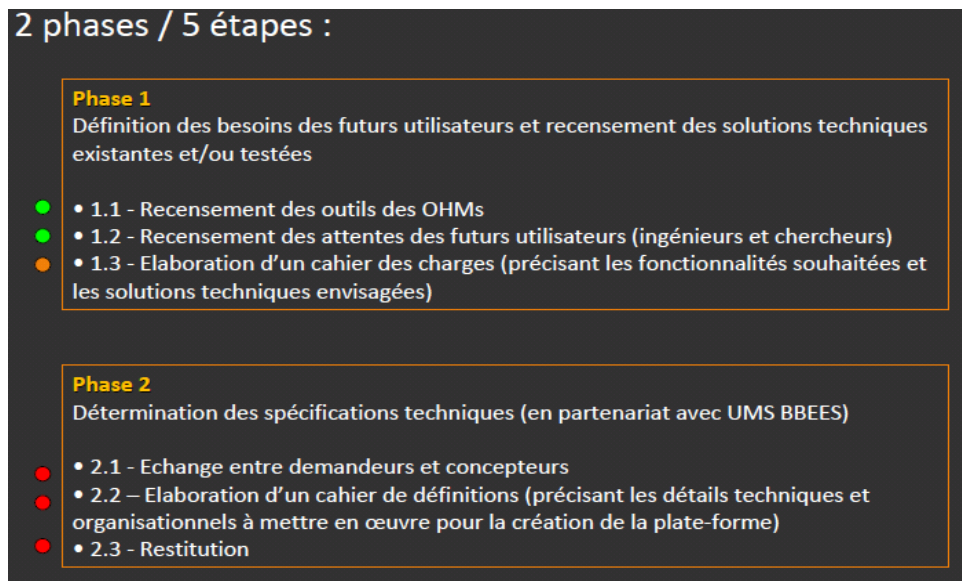


FIGURE 5 - EXEMPLES DE TABLEAUX DE BORD QUI ONT PERMIS DE SUIVRE L'AVANCEMENT DU PROJET

1.2. METHODES ET OUTILS

Les objectifs de la première phase du projet ont été atteints selon deux méthodes de consultation collective : 1) un recensement par formulaire et 2) des échanges lors de réunions de travail et lors de l'organisation d'un séminaire sur le thème des métadonnées.

Les personnes, qui ont participé à ce travail, ont été identifiées au cours du séminaire ROHM, organisé en mai 2012, à Lyon. Il s'agit essentiellement de personnel technique relevant des OHM (animateurs), du ROHM, et des équipes d'adossement des OHM (ingénieurs en informatique, géomaticiens, etc.). Il est à noter que la

volonté initialement affichée d'impliquer, dans cette démarche, les scientifiques participant aux études financées par les OHM a été difficile à mettre en œuvre.

I.2.1. RECENSEMENT PAR FORMULAIRE

Un questionnaire a été formalisé sous Google Documents pour le recensement des outils de gestion des métadonnées et des données existants dans les différents OHMs. Ce formulaire électronique a ensuite été diffusé auprès du groupe de travail. La grande majorité des OHM y a répondu ; seuls six des huit OHMs ont, à l'époque, été sondés : l'OHM Littoral Méditerranéen venait alors d'être créé et l'OHM Nunavik était toujours en cours de structuration.

I.2.2. ÉCHANGES ET REUNIONS

Au-delà des échanges informels et réguliers entre les pilotes du projet, l'ensemble des deux phases du projet ont permis l'organisation de 10 réunions de travail :

- **16/04/2012** : 1^{er} rencontre avec l'UMS BBESS, présentation du projet OGIIIS, discussion autour des modalités de coopération.
- **20/07/2012** : Discussion sur les attentes, les besoins et les contraintes des OHM dans le cadre du projet OGIIIS.
- **23/10/2012** : Séminaire de travail élargi autour des solutions techniques de gestion des métadonnées
- **28/07/2012** : Première présentation du projet au collectif OHM (séminaire de restitution de l'OHM BMP)
- **23/01/2013** : Deuxième réunion avec BBESS, présentation de la première version du cahier des charges
- **01/03/2013** : Troisième réunion avec BBESS, présentation des spécifications techniques à l'ensemble de l'équipe de l'UMS.
- **02 et 03/04/2013** : Conception en binôme du référentiel Acteurs/Projets.
- **09/04/2013** : Discussions sur l'interopérabilité des standards de description et le Web sémantique.
- **17/04/2013** : Réunion avec le collectif OHM : validation des choix techniques finaux relatifs à la plateforme
- **18/06/2013** : Présentation des résultats d'OGIIIS au cours du séminaire d'Aveiro.

Ces réunions ont fait l'objet d'un compte rendu. L'ensemble des propositions formulées à l'issue de la phase 1 a fait l'objet d'une validation par l'ensemble des directeurs d'OHM au cours d'une réunion qui s'est tenue à Toulouse, le 12 Novembre 2012.

II. RAPPEL DES PRINCIPAUX RESULTATS

Les principaux résultats de ce projet consistent : à avoir listé les outils existants de gestion des données présents dans les OHM d'une part, à avoir identifié au préalable les points de blocages potentiels et leviers d'action d'autre part et enfin, à avoir précisé les conditions organisationnelles, techniques et matérielles pour la mise en place de la plateforme. Ces résultats sont déjà en grande partie exposés dans la synthèse de faisabilité (cf. 0). Nous ne précisons ici que les éléments non préalablement détaillés. Un renvoi précisera les sections et items où sont explicités les principaux résultats. Les spécifications techniques de chacun des modules seront détaillées dans la section suivante.

II.1. RECENSEMENT DE L'EXISTANT : OUTILS, CONTRAINTES, ATTENTES

La première étape de l'étude de faisabilité a consisté à étudier le contexte de la question de la gestion de l'information scientifique au sein des OHM. Cette première démarche a eu pour buts de répondre aux questions suivantes :

- quelles sont les pratiques des OHM en matière de gestion de l'information scientifique ? Existe-t-il des outils spécifiques dans les OHM pour gérer les données ?
- Ces pratiques sont-elles satisfaisantes aux regards des ambitions du DRIHM ?
- Si tel n'est pas le cas, quelles en sont les raisons ? Existe-t-il des leviers qui permettraient de résoudre les problèmes rencontrés ?

Cette analyse du contexte actuel de la gestion des données dans les OHM a été menée à l'aide de questionnaires en ligne envoyés à chacun des OHM et via des réunions de travail où des représentants de chaque OHM étaient conviés. Il en ressort que la gestion par les OHM de leur information scientifique fait notamment appel aux outils suivants.

▪ Outils de gestion des métadonnées :

- [Cat'-BMP](#) (OHM BMP) :
- [EIVIS](#) (OHM VR) :

Ces deux outils collaboratifs de catalogage en ligne sont basés sur l'outil Open Source [GeoNetwork](#). Ils sont principalement utilisés pour la gestion des jeux de données géographique mais permettent la description d'autres ressources.

▪ Outils de gestion de la documentation :

- [EndNote](#) (logiciel propriétaire utilisé par l'OHM Oyapock),
- [Zotero](#) (outil Open Source, utilisé par l'OHM Oyapock et l'OHM VR),
- [Doc-BMP](#) (outil maison développé par l'OHM BMP pour gérer à la fois la littérature grise, la bibliographie scientifique, les vidéos).²⁷

▪ Outils de gestion de l'information spatiale :

- deux OHM (OHM BMP et OHM VR) ont mis en place des solutions de cartographie en ligne basées sur le *firmware* ArcGis Server, éditeur et diffuseur de services cartographiques.
- L'OHM VR est en cours de test d'un nouvel outil basé sur du code libre : [MapMint](#), Serveur cartographique développée par la société [Cartogenic](#).

▪ Outils de gestion de banque d'images ou photothèques :

- l'OHM Haut Vicdessos (HV) utilise une solution libre, développée à partir du logiciel Piwigo.

▪ Outils de gestion de contenu (CMS) :

Tous les OHM dispose d'un site ou d'une page web :

- basé sur SPIP pour les OHM [BMP](#), [OYA](#), [LM](#), [NUN](#)
- basé sur d'autres solutions : OHM [HV](#), [GMV](#), [EST](#), [VR](#).

Outre les applications en ligne, les métadonnées sont parfois saisies localement. Quelques outils prévus à cet effet ont été mentionnés lors des échanges :

- Morpho pour la saisie de métadonnées compatible avec la norme EML ;
- Exif Pilot, Exif Tool pour les descriptions de photos selon les standards correspondants ;
- les auxiliaires de saisies conçus pour INSPIRE (norme ISO19139) tels que le couple EXPIRE-ZAPHIRE ou les formulaires proposés par le CIGAL, le BRGM, etc.

²⁷ Cet outil n'est pas encore exploitable et il sera difficilement maintenu puisque ce type d'outil « maison » exige une évolution constante en lien avec les normes et les standards. Le choix de logiciels libres doit s'orienter de préférence vers ceux disposant d'une communauté active. Certains éléments de sa conception, les plus génériques, sont néanmoins utiles à OGIS.

En ce qui concerne le recensement des attentes, ces dernières ont été essentiellement exposées au cours de la réunion du 20/07/2012 par les différents participants. Elles sont partiellement exposées dans la sous-section 1.1.2.1. du présent document. Les fonctionnalités retenues (cf.1.II.1.) répondent à ces attentes.

Les obstacles à la mise en place de la plateforme sont cités dans la section 1.1.1.

II.2. DEFINITION DES SPECIFICATIONS GENERALES AU NIVEAU DE LA PLATEFORME

Ces éléments sont en partie développés dans la synthèse de faisabilité (chapitre 1.II).

II.2.1. HEBERGEMENT

Le service d'hébergement doit être dimensionné pour répondre au besoin d'un portail de données soumis à des requêtes et à des saisies concernant du contenu dynamique (enregistré dans des bases de données), mais aussi à des transferts de fichiers entrants et sortants. Cet hébergement sera vraisemblablement réparti sur différents serveurs, car OGIIS comporte plusieurs éléments de natures distinctes dont les exigences en termes de ressources et de contexte d'exploitation peuvent être différentes (cf. III.2.4.).

La recherche d'hébergement s'orientera vers des solutions prenant en charge ces technologies et/ou les prérequis spécifiques aux logiciels retenus. Sur ce dernier point, le présent document avance quelques pistes à prolonger à travers une veille technologique ad hoc. La compatibilité des offres disponibles avec les offres de services de la DSI du CNRS est donc un critère de choix. Cette dernière héberge déjà les sites internet de plusieurs OHM, ce qui constitue un antécédent favorable.

A terme, un transfert de l'hébergement vers l'IN2P3 est à prévoir (2016). Il sera encadré par l'UMS BBEES.

II.2.2. SUPERVISION ET MAINTENANCE

Les administrateurs doivent disposer d'une vue d'ensemble de l'infrastructure informatique et de la fréquentation.

Le logiciel Piwik²⁸ est déjà utilisé pour le suivi de la fréquentation de certains des web-services OHM et donne des résultats satisfaisants. La fréquentation ne représente néanmoins qu'un aspect de la supervision requise. Une étape de prospective logicielle serait utile pour identifier d'autres outils adéquats.

Les administrateurs doivent être avertis en cas d'obsolescence d'un enregistrement. Des scripts de vérification de la validité des fiches de métadonnées assureront cette fonction. Des requêtes automatiques régulières sur les URL, les URI, les adresses IP des serveurs, les adresses électroniques des contributeurs seront programmées avec une fréquence à définir selon les variables.

II.2.3. INTEGRATION WEB

OGIIS devra être doté d'une interface et d'un référencement lui assurant une visibilité dans le domaine de l'open access^{29, 30}. La mise en œuvre du protocole OAI-PMH^{31, 32} - déjà effective dans certains dispositifs de recherche comme le TGE Adonis - est particulièrement indiquée.

²⁸ <http://fr.piwik.org/>

²⁹ <http://openaccess.inist.fr/>

³⁰ <http://www.opendoar.org/>

Il serait avantageux, en termes de navigation, que les interfaces d'OGIIS tiennent compte de l'organisation et du graphisme des sites Web des OHM et de leur réseau, notamment ceux basés sur SPIP, en raison de son caractère conventionnel.

SPIP est proposé par le CNRS à ses équipes pour la création de sites web harmonisés. En tant que système de gestion de contenu, peut être mis en relation avec une ou plusieurs bases de données. Dans le cas des OHM, SPIP est le CMS majoritairement utilisé : il concerne 5 des sites internet (actifs ou en développement) sur les 9 que comptent les OHM.

D'autre part, la modularité de SPIP pourrait lui permettre de générer et/ou d'interpréter des flux de métadonnées, de sur-qualifier des données et d'afficher des web-services ; ces caractéristiques peuvent s'avérer intéressantes en termes de webmestrie. Incidemment, il peut servir d'instance de catalogage et de diffusion de contenus lorsqu'un catalogue spécialisé n'est pas souhaité, notamment si son installation est jugée trop complexe ou son administration trop chronophage.

III. SPECIFICATIONS TECHNIQUES DES DIFFERENTS MODULES D'OGIIS

Les spécifications techniques présentées dans ce chapitre sont basées sur les retours d'expériences organisés au cours de cette étude, sur l'expertise disponible au sein du réseau et ont été discutées avec l'UMS 3648 BBEEES. Ce chapitre dresse une ébauche des mécanismes qui devront opérer au sein d'OGIIS pour que celui-ci remplisse sa fonction d'auxiliaire du suivi scientifique.

III.1. CONSTRUCTION DU REFERENTIEL ACTEURS/PROJETS

La première des priorités consiste à établir un référencement des personnes et des actions de recherches rattachées au DRIIHM, aux OHM et à leur réseau. Les enjeux sont multiples : ils concernent à la fois la visibilité du dispositif de recherche, l'accessibilité de ses services en lignes, la sécurité et la fiabilité des informations véhiculées.

III.1.1. MISE EN PERSPECTIVE

Le référencement actuel des acteurs du ROHM constitue d'ores et déjà une base sur laquelle plusieurs propositions d'améliorations peuvent s'asseoir.

Le contrôle des accès aux bases de données d'OGIIS est un élément clé. Toute saisie dans OGIIS ne sera possible que pour les opérateurs enregistrés dans un annuaire et ayant reçu des coordonnées d'identification. Cette identification permettrait de se connecter à OGIIS en tant que tel mais serait aussi utilisable dans les applications Internet des OHM et de leur réseau telles que les Intranet de leurs sites web, les bases de données en ligne, les web-services cartographiques, les photothèques, les catalogues de métadonnées. Pour chaque personne, on souhaite connaître les informations suivantes : nom, prénom, organisation, numéro de voie, type de voie, nom de voie, complément d'adresse, code postal, ville, numéro de téléphone (format international), courriel(s), login, mot de passe. Ces coordonnées professionnelles doivent être maintenues à jour autant que faire se peut car elles sont nécessaires à la prise de contact et aux échanges d'informations comme cela sera expliqué dans les paragraphes suivants. Le super-administrateur sera en charge du contrôle des comptes d'utilisateurs.

³¹ http://fr.wikipedia.org/wiki/Open_Archives_Initiative_Protocol_for_Metadata_Harvesting

³² <http://www.openarchives.org/OAI/openarchivesprotocol.html>

Elles seront complétées par d'autres informations qui relèvent quant à elle de l'annuaire de compétences : fonction, spécialités, responsabilité éventuelle vis-à-vis des OHM, liens avec les projets OHM. Ces informations sont déjà enregistrées dans la base de données du site internet du réseau des OHM (www.ohm-inee.cnrs.fr). Cet annuaire de compétences est aujourd'hui consultable par les personnes internes au réseau. Il doit cependant connaître des évolutions significatives concernant en particulier l'organisation de sa base de données relationnelle et l'ergonomie de ses interfaces pour correspondre aux besoins exprimés dans le cadre d'OGIIS.

OGIIS n'a pas cependant comme vocation première d'assurer la maintenance d'un annuaire de personnes au-delà de ce qui touche directement à la traçabilité des données. Dans la mesure du possible, il serait souhaitable de rechercher une source d'information extérieure, suffisamment centralisée et conventionnelle pour mettre à jour les profils des utilisateurs aussi uniformément que possible. Avec l'essor du web collaboratif et des services en ligne, les utilisateurs d'internet voient se multiplier leurs espaces personnels, ainsi que les identifiants et mots de passe correspondants. En l'occurrence, les contributeurs potentiels au contenu d'OGIIS seront vraisemblablement identifiés auprès d'au moins un OHM, de leur(s) laboratoire(s) et de l'EPST de la recherche et de l'enseignement auquel ils appartiennent.

La possibilité de s'appuyer sur un annuaire institutionnel comme Labintel³³ ou sur des services déjà partagés au sein de l'enseignement supérieur et de la recherche comme EDUROAM³⁴ eut été intéressante de ce point de vue.

En l'absence d'une telle solution, la description des utilisateurs d'OGIIS se fera dans une instance centrale, commune à tous les OHM et dotée de mécanismes d'actualisation fiables. Il en va de la traçabilité et de la ré-utilisabilité des données car toutes deux exigent que les producteurs, les ayant-droits, les descripteurs de jeux de données puissent être identifiés sans ambiguïtés et contactés en cas de besoin.

Dans le domaine de la gestion des droits d'accès, le protocole LDAP (Lightweight Directory Access Protocol) est reconnu pour son efficacité. Celle-ci repose notamment sur optimisation des droits de lecture grâce à l'indexation d'entités organisées selon une arborescence stable. **OpenLDAP**³⁵ (libre), utilisé notamment par l'OSU Pythéas, constitue une solution logicielle libre et bien documentée pour l'application de ce protocole et son installation fait à ce titre partie des recommandations du présent document.

III.1.2. PROPOSITIONS POUR UNE NOUVELLE GESTION INTEGREE DES PROJETS DE L'APR DU ROHM

Les informations relatives aux projets sont pour la plupart saisies durant la période d'ouverture de chaque APR annuel. Depuis fin 2012, ces informations peuvent être saisies directement en ligne sur le site internet du réseau des OHM. Les informations sont enregistrées dans une base de données MySQL comme contenu du site internet en tant que réponse à un formulaire exportable au format CSV.

³³ <http://vhpora.dsi.cnrs.fr/labintel/>

³⁴ <http://www.eduroam.fr/>

³⁵ <http://fr.wikipedia.org/wiki/OpenLDAP>

Formulaire APR 2013

OHM concerné(s) (Obligatoire)

Veuillez cocher le ou les OHM concerné(s) par votre proposition.

- réseau des OHM
- OHM Bassin minier de Provence
- OHM Ojapock
- OHM Haut-Videssos
- OHM Tessékéré
- OHM Estarreja
- OHM Vallée du Rhône
- OHM Littoral méditerranéen

1. TITRE du projet

Titre complet (Obligatoire)

Sigle

FIGURE 6 - APERÇU DU FORMULAIRE DE SAISIE DES APR

Cette fonctionnalité doit encore connaître quelques améliorations. D'une part, les formulaires doivent évoluer vers plus d'ergonomie et de confort. D'autre part, parmi les attributs de chaque projet, certains relèvent de l'annuaire ou sont reliés à d'autres entités décrites dans d'autres modules, voire dans des ressources extérieures. Dans certains cas, les vocabulaires contrôlés seraient par ailleurs utiles pour assurer la cohérence des enregistrements.

Lors du prochain APR, toutes les candidatures devront être adressées par ce moyen. Les fonctionnalités répondant aux besoins des directeurs et des coordinateurs scientifiques d'OHM devront être intégrés dès la conception du modèle de données. Ils devront être consultés concernant notamment le rattachement des projets aux axes thématiques et les relations entre les différents projets : collaborations, échanges, prolongement de projet(s) antérieur(s).

Deux actions ont été entreprises à ce propos : la modélisation des informations (du modèle conceptuel de données au modèle physique de données) et la schématisation du processus global de gestion des projets de l'APR annuel du ROHM.

Le processus global de gestion des projets issus de l'APR annuel des OHM et de leur réseau a été défini en parallèle du modèle conceptuel de données (schéma ci-après). Cette double démarche a permis de vérifier le modèle et de mettre en exergue les dialogues de la future interface du système d'information. Ce processus se décline en trois phases temporelles distinctes :

1. **Réponse à l'APR ROHM de l'année n par les personnes intéressées** (en fin d'année n-1) : A l'issue de cette phase, chaque porteur de projet possède des login/mot de passe et le statut de chaque projet est « proposé ». Il est important de noter que toute saisie dans OGIIS ne sera possible que pour les opérateurs enregistrés dans le système et ayant reçu des coordonnées d'identification.
2. **Validation des projets de l'APR ROHM par les responsables des OHM**, après consultation et analyse des dossiers (en fin d'année n-1). Chaque directeur d'OHM accède à un formulaire de validation qui liste les projets déposés et en face de chaque projet, des cases à cocher permettent de retenir ou refuser le projet. A l'issue de cette phase, chaque projet a ainsi son statut qui passe de « proposé » à « retenu » ou « refusé » selon la décision prise. Le site web du ROHM donne la liste des lauréats. Les responsables des OHM sont libres de communiquer leurs avis aux porteurs de projets comme cela était le cas auparavant.

3. **Finalisation des projets par l'envoi du rapport d'activités par les porteurs des projets** de l'année n (début année n+1). Il ne s'agit pas ici de clôture administrative et financière, le système d'information n'ayant pas cette vocation. Le porteur de projet dépose le rapport d'activités via l'interface prévue à cet effet. Il peut également compléter certaines informations dans le formulaire « post-projet ». A l'issue de cette phase, les responsables des OHM concernés en sont informés et chaque projet reçoit le statut « clos ».

Indépendamment de ces trois étapes, une interface de consultation des projets retenus des années antérieures devra être prévue. Trois types d'accès seront à prévoir :

- un accès grand public pour la consultation d'informations générales sur le contenu scientifique des projets ;
- Un accès réservé aux acteurs (login/mot de passe) leur donnant la possibilité de consulter tous les projets et les coordonnées des acteurs ;
- Un accès réservé aux responsables des OHMs (directeurs, ingénieurs support), permettant par exemple de donner des informations générales sur les projets et de fournir certaines statistiques (nombre de projets, de participants, de thématiques abordées, coût moyen d'un projet, etc.).

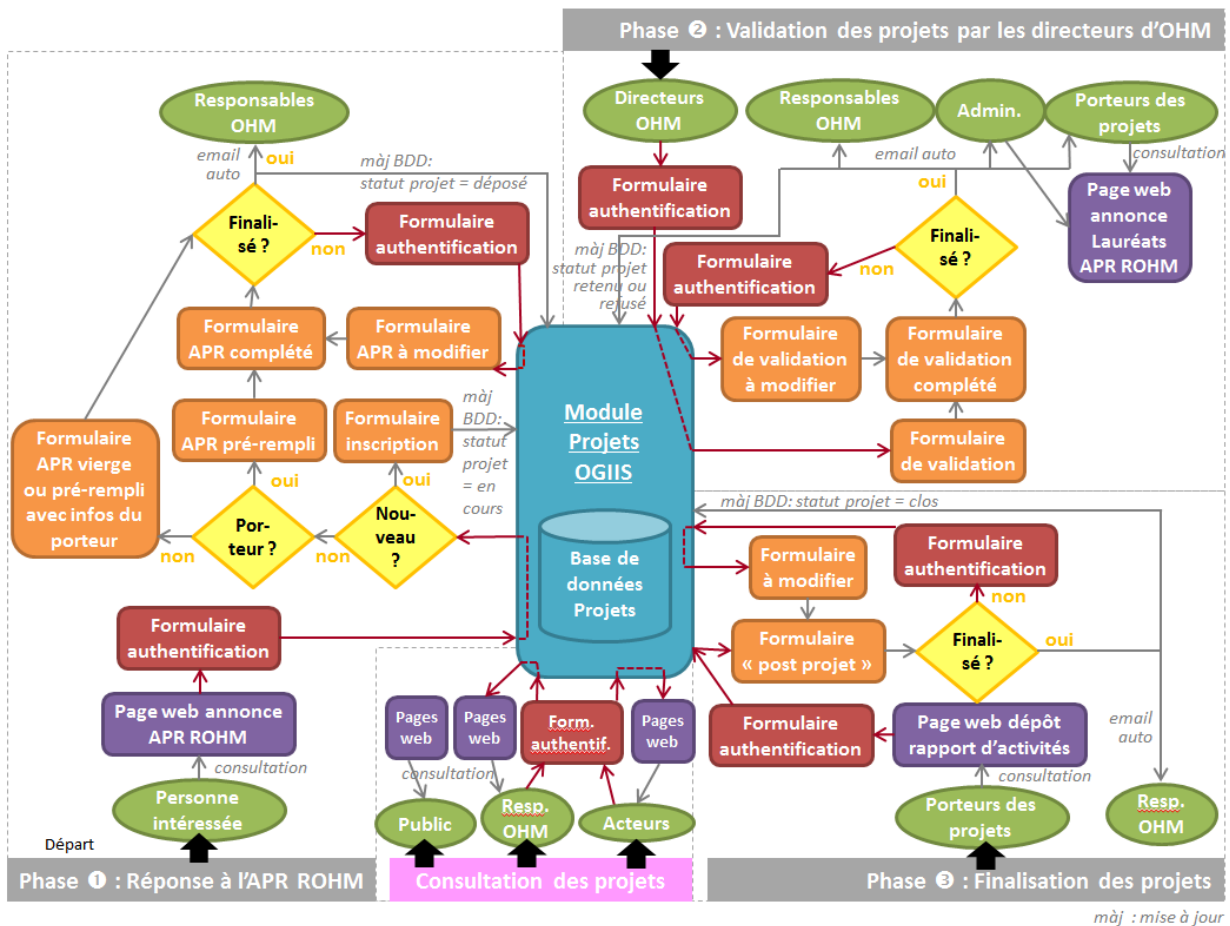


FIGURE 7- PROCESSUS DE GESTION DES PROJETS DE L'APR ANNUEL DU ROHM

S'agissant de l'interface, les maquettes des principales pages web identifiées pour guider les personnes dans le processus sont présentées en annexe :

- formulaire d'inscription à OGII (annexe 5.02).
- Formulaire de saisie en ligne des projets pour l'APR annuel du ROHM (annexe 5.03).
- Interface de validation des projets par le responsable OHM (annexe 5.04).
- Formulaire « post projet » de saisie en ligne des rapports d'activités (annexe 5.05).

Par ailleurs, dans le but d'harmoniser autant que possible le contenu du rapport d'activités, il sera prévu dans l'interface la possibilité de télécharger une **trame de rapport d'activités**. Il reste toutefois à valider son plan.

III.1.3. DETAIL DE LA MODELISATION DE LA BASE DE DONNEES DU REFERENTIEL ACTEURS/PROJETS

Un modèle logique de données (MLD) relatif aux projets de recherche financés ou cofinancés par l'APR annuel du ROHM ont été modélisées (annexe 5.06). Il est inspiré sur l'analyse des trois outils existants suivants :

- le site web du ROHM sous SPIP permettant de gérer les données issues du formulaire de saisie en ligne des projets pour l'APR annuel (stockage en base de données sous MySQL) ;
- La base de données « projets OHM Haut-Videssos » (PostgreSQL/PostGIS et interface web cartographique avec les bibliothèques de fonctions Javascript GeoExt, Openlayers et ExtJS), en particulier son modèle conceptuel de données ;
- Le projet de base documentaire « doc-BMP » (Zotero/Excel), en particulier son modèle logique de données.

Le modèle physique de données (MPD) apporte principalement les informations suivantes :

- La table PROJET gère les informations spécifiques à chaque projet. Deux vagues de saisie permettront de la compléter :
 - via le formulaire de saisie en ligne des projets pour l'APR annuel (formulaire existant créé pour l'APR 2013 révisé suite à la modélisation) ;
 - via le futur formulaire « post-projet » permettant le dépôt en ligne du rapport d'activités et la saisie de quelques informations complémentaires.
- Un projet peut concerner un à plusieurs OHM et/ou au réseau ROHM.
- Un projet peut être décrit par un ou plusieurs domaines scientifiques. Le terme de « spécialité » est utilisé pour désigner ces ensembles supposés homogènes de compétences, de méthodes et d'objets d'études. Il y a deux options exclusives possibles pour la gestion des spécialités :
 - un champ texte avec la liste des spécialités séparées par des virgules ;
 - un répertoire des spécialités lié à une ressource externe par des interactions qui restent à définir : flux XML (importé dans une table intermédiaire ou interrogé à la source à chaque saisie), modération des saisies d'utilisateurs (gestion des néologismes et des doublons), validation empirique et mécanismes d'auto-complétion.
- Un projet peut être décrit par un ou plusieurs mots-clés. Ces mots-clés ont une valeur sémantique sensiblement différente des spécialités. Ils décrivent le sujet lui-même des investigations plus que le traitement disciplinaire qui lui est accordé dans l'étude. Cet élément diffère cependant peu en termes d'alternatives fonctionnelles dans le système d'information :
 - un champ texte avec la liste des mots clé séparés par des virgules ;
 - une table répertoriant les spécialités via une ressource externe par des interactions qui restent à définir avec les mêmes interrogations que pour les spécialités.
- Un projet peut être relié à un à plusieurs axes de recherche, eux-mêmes reliés à un OHM distinct ou au ROHM. La relation physique est à l'image de la pratique liée à la gouvernance du dispositif de recherche : un axe est défini par et pour un et un seul OHM. Ce module devra permettre des études statistiques mais pas nécessairement les exécuter : les opérateurs autorisés (référents OHM, membres du CoPil d'OGIIS) peuvent effectuer les traitements sur des fichiers exportés (csv par exemple).

- Un projet aura un et seul rapport d'activités associé appelé ici « ressource ». Toutefois, dans la modélisation il a été choisi d'autoriser le fait qu'un projet puisse être associé à une ou plusieurs ressources. Cette version évolutive pourra à terme gérer d'autres ressources comme des fichiers de données, images, publications, etc.. Dans un premier temps, il est prévu de créer un formulaire permettant le dépôt en ligne du rapport d'activités et la saisie d'informations complémentaires. La possibilité de classer chaque ressource dans une « catégorie » sera aussi offerte : au départ il n'y aura qu'une seule catégorie pour les rapports d'activités, ensuite une typologie exclusive et exhaustive sera établie pour le périmètre des ressources potentiellement concernées par l'outil. La gestion des métadonnées associées sera différenciée selon les catégories. Ces dernières devront être clairement listées et validées (module Métadonnées).

- Un projet peut être lié à un ou plusieurs autres projets. Deux possibilités non exclusives pourraient être envisagées.
 - Dans le cas d'un projet prolongeant un projet antérieur (ex : DYVAH 2010, DYVAH 2011 et DYVAH 2012 de l'OHM Haut-Videssos), l'information pourrait être saisie au moment de la réponse à l'APR dans le formulaire dédié, et ce, par chaque porteur de projet ;
 - Dans le cas d'un projet en lien avec d'autres projets de noms différents et de la même année ou non (ex : EESIPRAM, MODHYC, P2AS, EMPAL, DYVAH 2010 et OPA de l'OHM Haut-Videssos), la saisie de ces liens ne pourrait se faire que par l'un des responsables de l'OHM.

- Un projet est associé à un ou plusieurs acteurs appelés « individus ». Chaque individu peut être qualifié soit de responsable soit de participant (rôle). Le fait que la liste des participants fournie en réponse à l'APR ne soit pas exhaustive constitue une limite : le candidat ne connaît pas toujours l'ensemble des personnes qui, à terme, auront contribué au projet en raison notamment des rencontres de mi-parcours, des recrutements saisonniers de stagiaires, etc... Le modèle proposé ici gère les informations suivantes.
 - Un acteur est en lien avec un ou plusieurs OHM et/ou le réseau ROHM.
 - Un acteur a une ou plusieurs spécialités (compétences). Pour la gestion de ces spécialités, les deux options présentées plus haut sont également valables ici.
 - Un acteur peut être rattaché à zéro ou un seul organisme (EPST, EPIC, société privée, association, etc.) à une date donnée. Si cet acteur participe à un nouveau projet et qu'il a changé d'organisme, cela sera pris en compte.
 - Un acteur peut être rattaché à zéro ou un seul laboratoire (ce dernier étant associé à un ou plusieurs organismes) à une date donnée. Si cet acteur participe à un nouveau projet et qu'il a changé de laboratoire, cela sera **pris en compte**. Les informations sur le laboratoire sont minimales. Pour en savoir plus sur le laboratoire, il suffira de se connecter sur les annuaires nationaux (ex : [LabIntel](#) pour le CNRS, <http://annuaire.inra.fr> pour l'INRA).
 - Un acteur peut avoir zéro ou une seule adresse à une date donnée. Si cet acteur participe à un nouveau projet et qu'il a changé d'adresse, cela sera géré.

- Les accès aux différents services dont OGIIS sont définis pour une ou plusieurs personnes. En revanche, tous les participants n'auront pas obligatoirement un login/mot de passe. A minima il s'agit des responsables des OHM, ingénieurs support des OHM, personnes en lien avec le ROHM, ainsi que les porteurs de projets. La table SERVICE répertorie les différents serveurs en lien direct avec les OHM : ELVIS, Cat-BMP, etc.

- Les informations fournies par les inscriptions et les propositions de recherche devront être réutilisées comme variables d'environnement pour faciliter la saisie de formulaire ultérieurs.

III.2. CONSTRUCTION DU REFERENTIEL METADONNEES

L'intérêt du catalogage au sein du réseau des OHM a déjà été exprimé et des exemples existent déjà en ce sens (cf. 1.II.1.1.). La directive INSPIRE et d'autres initiatives institutionnelles (Open Data, modernisation des services de l'état, sciences collaboratives, réponses aux demandes sociétales) créent un contexte favorable dans lequel l'agilité repose sur l'assimilation et l'harmonisation de pratiques. Ces pratiques couvrent de multiples aspects de la vie numérique des données : conceptualisation, indexation, etc. L'ouverture aux communautés voisines est elle aussi un facteur d'interopérabilité.

Il est à parier que les retombées en termes de visibilité pour les OHM serviront non seulement leurs desseins scientifiques mais aussi la viabilité de leurs outils en manifestant leur présence dans des réseaux d'expertises où s'élaborent les théories et les techniques du traitement de l'information.

III.2.1. FORMULAIRES DE SAISIE

La première étape de saisie consiste pour l'utilisateur à choisir un modèle de description parmi ceux déjà présents dans l'outil. Par défaut, c'est le Dublin Core qui sera proposé dans le menu déroulant présentant les différents modèles disponibles. Une fois le choix effectué la possibilité lui sera offerte de téléverser un fichier XML ou de saisir lui-même les informations demandées qui pourront éventuellement être pré-saisies en fonction de l'environnement de saisie.

L'ajout progressif de nouveaux modèles et leur encapsulation dans le XSD (XML Schema Definition)³⁶ du schéma RDF (Resource Description Framework)³⁷ est à prévoir. La procédure pour de tels ajouts reste cependant à définir. En admettant qu'elle soit ouverte aux contributeurs, deux options « types » peuvent a priori être envisagées.

- **Option 1** : L'utilisateur, dans un mode « expert », téléverse une trame de description RDF dont la syntaxe - notamment le balisage - est vérifiée automatiquement³⁸. Si la trame est valide, le téléversement du fichier de métadonnées est proposé. Si les métadonnées respectent le modèle fourni, alors le jeu de données est soumis à la publication comme s'il avait été complètement saisi.
- **Option 2** : L'utilisateur propose son modèle par courriel en pièce jointe sous la forme de fichier XML en expliquant sa construction, son intérêt et son articulation avec les modèles déjà implémentés.

L'administrateur doit alors intégrer ce modèle à ceux existants après vérification de la compatibilité avec le cadre existant fourni par l'application.

En fonction du modèle choisi, le squelette assurera l'affichage du formulaire correspondant. Le respect des conventions de saisie sera contrôlé *a posteriori* par des mécanismes applicatifs lorsque l'affichage n'aura pas permis de prévenir les éventuelles erreurs.

La conception de ces formulaires de saisies consiste, dans un premier temps, à en établir les variables, les contraintes et les conventions en tenant compte de l'exigence d'ergonomie. Une partie de ce travail a déjà été effectué dans le cadre du projet de base de données documentaire de l'OHM Bassin minier de Provence.

³⁶ http://www.w3schools.com/schema/schema_intro.asp

³⁷ <http://www.w3.org/RDF/>

³⁸ A titre d'exemple, voir <http://validator.w3.org/>.

III.2.2. CONSULTATION

Les interfaces de consultation présenteront des mécanismes de requêtes déjà utilisés dans des applications connues telles que GeoNetwork.

- **Requêtes textuelles** : un champ de saisie libre permet de rechercher une valeur (chaîne de caractère ou nombre) dans les enregistrements de la base de données.
- **Requêtes spatiales** : une fenêtre cartographique permet de définir un périmètre d'intérêt au sein duquel tous les jeux de données ayant une emprise seront sélectionnés.
- **Classement** : l'utilisateur pourra aussi choisir dans différentes listes de valeurs classées de lieux, de mots-clefs, d'équipes de recherche, d'auteurs ou encore d'année de publication.

Par défaut, une fiche de métadonnées s'affiche d'après la trame description avec laquelle elle a été enregistrée dans OGIIS. L'utilisateur doit, au cours de ses recherches, pouvoir sélectionner les résultats qui l'intéressent et les exporter au format XML selon le standard de son choix, parmi ceux implémentés.

- **Métadonnées embarquées**

La norme XMP s'appuie sur quelques standards de base, dont le Dublin Core, et présente l'avantage d'être extensible à d'autres standards de métadonnées. Elle permet d'incorporer les métadonnées aux fichiers de données de telle sorte qu'elles soient consultables en tant que propriétés du document.

La correspondance entre le contenu des descriptions et le schéma RDF permet l'interprétation et la manipulation des métadonnées par les machines. L'encodage automatique de métadonnées embarquées dans les fichiers de données selon la norme XMP pourrait ainsi être assuré par un script exécutable à chaque insertion de nouvelles données dans la base. Cette procédure d'intégration est déjà formalisée pour un certain nombre de formats : JPEG, TIFF, GIF, PNG, HTML, PDF, XML/SVG, AI, EPS.

III.2.3. GESTION DYNAMIQUE DES METADONNEES

Les métadonnées doivent être gérées de manière dynamique pour faciliter leur actualisation. Des procédures automatiques doivent, dans la mesure du possible, contrôler la validité des données ou a minima informer les administrateurs des opérations de maintenance à réaliser. Ces mécanismes de supervision se réfèrent, entre autres, à des sources d'informations extérieures à l'outil proprement dit. Ils concernent notamment :

- le respect des standards de description (triplet RDF, modèle XSD),
- le contrôle des vocabulaires (thesaurii),
- la validité des adresses (URL, IP, URI, résolveurs),
- le bon fonctionnement de l'application (scripts),
- l'identification des opérateurs (annuaire),
- le cycle de vie des données (durée des projets, des droits, etc...).

Le W3C recommande de baser les descriptions de ressources sur des trames de description de ressource RDF (Ressource Description Framework³⁹). OGIIS devra concevoir et formaliser son propre modèle de description de ressource intégrant tous les standards retenus, ne serait-ce que pour être en mesure d'intégrer la diversité des descriptions fournies.

Le contributeur doit être, le cas échéant, en mesure d'associer aux métadonnées qu'il fournit la trame de description associée, si celle-ci n'est pas prévue.

39 <http://www.w3.org/RDF/>

De nombreux éléments de métadonnées requièrent l'usage de vocabulaires contrôlés. Cela vaut notamment pour les mots-clefs où le recours à des thesaurii internationaux assure, pour partie, le multilinguisme, incontournable dans le cas d'un dispositif de recherche international, tel que le réseau des OHM. Dans une certaine mesure, on peut également penser que les ontologies serviraient la pratique de l'interdisciplinarité. D'une part, elles peuvent jouer un rôle didactique. D'autre part, elles fourniront un cadre sémantique aux descriptions saisies et faciliteront les interprétations humaines et automatisées auxquelles elles sont destinées.

C'est également valable pour les noms de lieux, les types de ressources (MIME), les formats et volumes caractéristiques de chaque ressource. D'après les bonnes pratiques suggérées par les éditeurs de standards de description et/ou déjà en vigueur dans les OHM, un certain nombre de vocabulaires est *a priori* indiqué et disponible. Plusieurs thesaurii pourront être proposés pour un même champ. Leur liste doit être extensible : les contributeurs pourront en suggérer de nouveaux auprès des administrateurs d'OGIIS. La saisie libre sera toujours accordée aux contributeurs. Elle pourrait dans certains cas être conditionnée par le renseignement d'au moins un élément relevant d'un vocabulaire contrôlé.

La plupart des champs devront aussi faire l'objet de conventions de saisie claires, pour faciliter l'automatisation des traitements: dates, noms, droits liés à la donnée, titres, systèmes de coordonnées, et d'une manière générale tous les éléments dont l'interprétation peut varier en fonction de la manière dont ils sont renseignés. Les formulaires de saisie joueront un rôle essentiel dans le respect de ces conventions en s'appuyant sur la signalisation (menus d'aides contextuels), l'ergonomie (menus déroulants, séparation de champs, pré-formatage des variables) et de contrôles intégrés. Cet aspect sera détaillé dans la partie « Interfaces ».

La gestion unifiée doit prévenir la formation de doublons, la persistance de données obsolètes dans les services les moins fréquemment actualisés et assurerait ainsi la validité des informations fournies. OGIIS en tant que référentiel de données (Master Data⁴⁰) constituera ainsi le premier maillon d'une chaîne de systèmes d'agilité de gestion (ACMS⁴¹).

Le catalogage des métadonnées comprendra les attributs intellectuels et techniques relatifs à la donnée. L'indexation des enregistrements assurera le lien entre les différentes classes d'attributs réparties selon le schéma relationnel. Ce schéma relationnel, en cours d'élaboration sur la base de travaux existants, devra assurer l'identification uniforme des ressources et permettre la surqualification.

Les attributs devront être organisés de la manière la plus rigoureuse possible au sein de ce schéma. A titre d'exemple, il est plus sûr de concaténer deux champs que de distinguer automatiquement des séparateurs formels dans des enregistrements saisis par des opérateurs différents.

Des identifiants de ressources uniformes persistants (URI) doivent être attribués chaque fois que possible, en plus d'une éventuelle URL. Les URI suivants peuvent être cités à titre d'exemple : ISSN, ISBN, ISRN⁴², URN:NBN⁴³, PURL⁴⁴, handle⁴⁵, DOI⁴⁶, ARK, LSID⁴⁷, ISGN⁴⁸.

⁴⁰ <http://dl.orchestranetworks.com/sita/2012/IntroductionToMDMPart1.pdf>

⁴¹ <http://www.sustainablearchitecture.com/acms.html>

⁴² http://www.bnf.fr/fr/professionnels/s_informer_autres_numeros/a.isrn_autres_numeros.html
http://www.adbs.fr/isrn-numero-international-normalise-des-rapports--17515.htm?RH=OUTILS_VOC

⁴³ <http://www.nationallibrary.fi/publishers/urn.html>

⁴⁴ <http://purl.oclc.org/docs/index.html>

⁴⁵ <http://www.handle.net/>

⁴⁶ <http://www.crossref.org/>

⁴⁷ <http://en.wikipedia.org/wiki/LSID>

Outre la base de données proprement dite, ce module de méta-description comprend les triplets RDF garant de l'interopérabilité et la couche applicative constituée des traitements automatiques applicables aux métadonnées enregistrées.

III.3. CONSTRUCTION DU MODULE D'ARCHIVAGE

En tant que dispositif d'observation à long terme, les OHM doivent bénéficier d'un niveau de service élevé pour la sauvegarde de leurs données. Une planification sur une durée de l'ordre supérieur à la décennie est nécessaire pour assurer la continuité de l'information produite et conservée dans ce cadre.

Du point de vue des techniques d'archivages, l'interopérabilité est aussi une question de supports physiques. Les durées envisagées pour l'archivage à long terme posent la question de la durée de vie des matériaux et des technologies. Les formats d'archivage sont choisis pour leur pérennité. C'est l'une des raisons pour laquelle les standards ouverts dont la syntaxe et la sémantique resteront interprétables doivent être privilégiés dès la production. Dans le cas contraire, la conversion sera un préalable impératif à l'archivage.⁴⁹

On souhaite un archivage à long terme sans perte (pas de sauvegarde glissante). Le recours à d'éventuels algorithmes de compression doit tenir compte de cette contrainte, étant entendu qu'il s'agit d'éviter de quelconques dégradations irréversibles des données archivées. En revanche, un temps de réponse plus long que pour une application courante serait a priori acceptable pour la plupart des utilisateurs.

Les protocoles d'archivages sont normalisés et la loi fournit à bien des égards un cadre assez strict à ce sujet, notamment à la durée des droits. C'est en connaissance de ces éléments qu'OGIIS devra définir les scénarii d'archivage correspondant à ses objectifs et le niveau de dématérialisation relatif à l'administration de ce service.

Les interfaces d'archivage doivent permettre l'envoi et la récupération de fichiers par les administrateurs des données relevant d'OGIIS sur demandes des instances responsables de ces données (administrateurs OHM).

La volumétrie et la répartition des droits d'accès par opérateur (visiteur, contributeur, administrateur) sont des points cruciaux sur lesquels les coordinateurs et directeurs d'OHM seront amenés à se prononcer.

L'archivage requiert un savoir-faire particulier et une fiabilité sur le long terme. Des services spécialisés et déjà engagés dans la sauvegarde des données publiques, voire des données publiques de la recherche, seront vraisemblablement les plus capables de répondre à ce besoin.

III.4. CONSTRUCTION DU MODULE GEO-VISUALISATION

Le module cartographique d'OGIIS sera composé d'une voire plusieurs bases de données et d'un éditeur de service. Des interfaces permettront à l'utilisateur de visualiser le résultat d'interrogations prédéterminées de ce module sous forme de cartographies interactives.

La norme INSPIRE constitue le standard indiqué pour la description des données géographiques. L'emprise des jeux de données est l'un des éléments requis par ce standard. Cette emprise pourra être visualisée sur des fonds cartographiques de repérage et faire l'objet de requêtes spatiales.

⁴⁸ http://en.wikipedia.org/wiki/International_Geo_Sample_Number

⁴⁹ http://references.modernisation.gouv.fr/sites/default/files/RGI_Version1%200_1.pdf

Le contenu du jeu de données peut être soumis à des contraintes de propriétés et de confidentialité : sa publication doit donc faire l'objet des contrôles techniques et administratifs préalables répondant à ces contraintes.

D'après l'état des lieux, les cartographies en ligne proposées par les OHM sont, de manière prédominante, diffusées au format WMS en utilisant le logiciel propriétaire ArcGIS Server. Les possibilités offertes par les logiciels opensource tels que Geoserver, PostGIS ou encore des solutions composites comme geOrchestra doivent être examinées avec attention car elles offrent a priori de meilleures perspectives en termes d'interopérabilité et de compatibilité descendante. La cohabitation à court terme de plusieurs solutions logicielles, libres ou propriétaires, doit être envisagée.

L'éditeur de services cartographiques requiert un hébergement sur une machine dédiée. D'une part, les spécificités d'une telle application l'exigent en grande partie. C'est ce que tend à démontrer l'expérience de l'OHM Bassin minier de Provence dans ce domaine⁵⁰. D'autre part, plusieurs éditeurs de services cartographiques pourront être amenés à coexister puisque certains OHM disposent déjà du leur. Il est souhaitable que de tels outils aient au sein du réseau des OHM une certaine autonomie. Les services publiés n'en seront pas moins ouverts car une attention particulière sera portée à l'interopérabilité des formats de diffusion.

III.5. CONSTRUCTION DU MODULE STOCKAGE/PARTAGE

III.5.1. RECEPTION DES DONNEES

La production des données relève du processus métier de gestion de données défini au niveau de chaque OHM. L'intégration des données au référentiel que constitue OGIIIS relève quant à elle du Réseau des OHM en tant que point de convergence.

Les formats ouverts et/ou représentant des standards de fait seront privilégiés, notamment dans la mesure où ils permettent l'intégration de métadonnées embarquées.

Afin d'assurer la lisibilité du fichier de données fourni, la saisie du type MIME dans les métadonnées serait obligatoire pour chaque fichier déposé.

L'harmonisation des jeux de données réceptionnés devra suivre un mode opératoire prédéfini dont l'élaboration fait partie intégrante de la conception d'OGIIIS.

III.5.2. INTEGRATION

Deux types de contributeurs sont susceptibles de téléverser des données :

- les producteurs eux-mêmes, c'est-à-dire les personnes liées à un projet de recherche (accepté lors d'un APR) et identifiées au sein du système d'information en tant que telles ;
- les responsables d'OHM (directeurs et coordinateurs) après avoir eux-mêmes accusé réception des livrables d'un ou de plusieurs projets arrivés à échéance.

Les producteurs de données, i.e. participants des projets, auront par défaut des droits d'accès au système d'information plus restreints que les responsables d'OHM. Ces derniers peuvent néanmoins être eux-mêmes producteurs de données.

⁵⁰ Contacts : J. Fleury (fleury@cerege.fr), A. Rotereau (rotereau@cerege.fr)

Les responsables d'OHM devront participer à l'intégration des données avec le support et sous la supervision de l'administration technique centrale d'OGIIS.

Cette intégration comprend les préparations et les vérifications destinées à harmoniser et à rehausser le niveau de référencement et d'interopérabilité des données.

III.5.3. INTERFACES

Le contributeur souhaitant téléverser une donnée devra, tout d'abord, saisir ou appeler la fiche de métadonnée correspondante. Cela implique des liens entrants et sortants avec les interfaces de consultation et de requêtes. Quelques opérations accompagneront le téléversement lui-même. La solution retenue devra les prendre en charge.

Proposer d'uploader du fichier de données immédiatement après l'enregistrement des métadonnées incitera au dépôt dans la foulée. Cette interface d'upload devra présenter une continuité maximale avec le formulaire de saisie des métadonnées pour ne pas dérouter le contributeur.

Le contributeur étant invité à préciser le cycle de vie⁵¹ et les droits attachés à la donnée, le dépôt n'entraîne pas systématiquement la publication. Seul un administrateur peut procéder à la publication d'une donnée à condition d'avoir préalablement recueilli les autorisations de l'auteur et/ou du détenteur des droits dans tous les cas où elles sont requises par la loi.

A l'affichage, pour une ergonomie optimale, l'utilisateur doit pouvoir naviguer entre la donnée et la métadonnée très rapidement (trois clics maximum).

III.5.4. GESTION DES FLUX

Les données disponibles fournissent des estimations de volume de l'ordre de 1To par OHM et par an (fourchette haute). Une procédure d'intégration rigoureuse est susceptible de diminuer la taille des fichiers. Il est cependant préférable de surévaluer la volumétrie car la collecte systématique des données se met progressivement en place au sein des OHM, d'où une montée en charge prévisible et potentiellement importante du stockage de fichiers.

III.5.5. HEBERGEMENT

Les serveurs d'hébergement doivent assurer un bon niveau de service et permettre la montée en charge des web-services fournis dans le cadre d'OGIIS. Les puissances de calculs et les débits des connexions internet doivent être dimensionnés en conséquence.

La fréquentation actuelle du site ohm-inee.cnrs.fr est de l'ordre de 500 visites par mois. D'après les données de 2011-2012, ce chiffre peut doubler voire tripler lors des appels à propositions de recherches annuelles.

Avec 80 projets retenus lors de l'appel à proposition de recherches 2013 la répartition moyenne par OHM est proche de 11. En 2014, 8 OHM pourront recevoir de nouvelles propositions de recherches. Le nombre de participants par projet varie grosso modo de 1 à 10. 26 doctorants et post-doctorants intégreront le DRIIHM

⁵¹ Michener, W. K. & Jones, M. B. Ecoinformatics: supporting ecology as a data-intensive science. *Trends in Ecology & Evolution* **27**, 85–93 (2012).

entre 2013 et 2021. A ces contributeurs potentiels s'ajoutent le personnel permanent affecté aux OHM et les partenaires extérieurs. Dans ces conditions, le nombre d'utilisateurs potentiels d'OGIIS semble suffisamment élevé pour que la question du dimensionnement de l'accès au service soit anticipée. D'autant plus que le nombre de requêtes peut s'avérer particulièrement élevé sur des périodes restreintes de l'année comme les lancements d'APR ou à l'approche des dates limites de restitution des livrables.

L'hébergement des fichiers de données proprement dit doit être séparé de l'hébergement de la base de métadonnées. Tout d'abord, la volumétrie de cet espace de stockage sera beaucoup plus importante car un seul fichier de données pèse en général plus lourd que l'enregistrement d'une fiche de description dans une base de métadonnées de type SQL - voire que la totalité d'une telle base. De plus, le système de gestion et les protocoles seront vraisemblablement différents, comme on le verra plus tard.

Cet espace de stockage sera accessible par des utilisateurs et des services distants. Cet accès sera contrôlé par une procédure d'identification. La publication des données par des web-services devra être soumise à des vérifications automatiques interrogeant les métadonnées. Les administrateurs OGIIS, voire les responsables d'OHM, pourront accéder à cet espace de stockage pour y déposer les données dûment préparées.

CONCLUSION & PERSPECTIVES

Le projet OGIIS dans son état actuel d'avancement a éclairci le chemin vers une infrastructure commune de données propres aux OHM et ouvertes vers leurs partenaires et collaborateurs. Il a permis d'identifier et de solliciter différents relais utiles pour un tel parcours. Il contribue d'ores et déjà à une meilleure connaissance d'un contexte institutionnel et réglementaire contraignant mais favorable à la réalisation des objectifs propres aux OHM, dans la mesure où ses exigences jalonnent et encouragent les efforts fournis en termes d'organisation des données.

OBSERVATIONS ET INFORMATIONS DANS LE SYSTEME OHM

L'activité des OHM, centrée sur l'observation des relations hommes-milieus avec des débouchés scientifiques et opérationnels à court, moyen et long termes les place d'emblée au cœur des problématiques d'interopérabilité et de durabilité des systèmes d'informations. Elle est concernée par les évolutions technologiques et réglementaires qui affectent la société dans son ensemble et plus particulièrement l'action publique et la recherche.

Formalisée par l'existence du Réseau des OHM, cette communauté d'objectifs est propice à l'apparition d'une communauté de pratiques en adéquation avec leur pluralité thématique et géographique. Cette position peut s'avérer stratégique pour développer un savoir-faire privilégiant la mise en valeur des connaissances interdisciplinaires et intégrant la complexité des systèmes de recherche déployés pour leur production. Enfin, les ressources allouées à ces efforts seront d'autant plus justifiées qu'elles représentent une mutualisation et une rationalisation. Les gains en fiabilité, en visibilité et en robustesse à espérer pour la chaîne de production de l'information constituent des arguments forts.

Pour y parvenir, les solutions techniques sont nombreuses et variées. L'ergonomie, l'indexation et la normalisation sémantique figurent en tête de liste des recommandations à appliquer pour optimiser le cycle de vie des données. Les réflexions théoriques et méthodologiques dans ce domaine gagneront à être maintenues et développées en tâches de fond à travers une veille technologique, voire par des projets de recherche dédiés. Le DRIIHM peut ainsi se donner la possibilité d'augmenter et d'homogénéiser le niveau de maîtrise des technologies disponibles et adaptées à sa programmation scientifique.

L'interopérabilité implique des conceptions visant à ce que le plus grand nombre de données puisse être interprétable aussi bien par des humains que des machines indépendamment de leur instanciation. L'enjeu est ici non seulement technique mais aussi directement scientifique puisque le croisement de données produites en des temps et des lieux différents par des personnes distinctes offre des perspectives de recherche démultipliées.

OGIIS : VERS UN NOUVEAU TYPE DE RELATIONS HUMAINS-OHM ?

Une attention particulière devra être accordée aux moyens humains mobilisables sur ces questions en son propre sein mais aussi à l'extérieur et auprès de ses partenaires et collaborateurs. D'une part, le pilotage de l'outil lui-même demande une certaine coordination. D'autre part, les développements et la maintenance informatiques peuvent eux-mêmes demander des compétences techniques diverses qu'il s'agit de réunir. Enfin, Il faut prendre en compte l'utilisateur final à travers une réflexion sur l'ergonomie de l'outil et son interdépendance avec les web-services existants ou à venir. C'est en prenant en compte les comportements de

navigation et les flux de données que pourra être construit un système de gestion de l'information scientifique en ligne pleinement intégré.

Dans un ouvrage intitulé « Système d'information et développement durable » [Tassin 2010] on peut lire les considérations suivantes.

« Nous savons que la plupart des décisions prises par les entreprises ainsi que leurs processus opérationnels s'appuient sur des logiciels qui incarnent leurs modes de fonctionnement. Lorsque cette incarnation se fige, c'est toute l'entreprise qui est paralysée.

[...]

Ces solutions sont écrites dans des langages informatiques que les responsables et les utilisateurs métiers ne maîtrisent pas. Avec le temps, les documentations de ces systèmes complexes deviennent obsolètes et la connaissance se perd, même pour les informaticiens. »

Chacun pourra juger de la pertinence de ces propos en se référant à sa propre expérience. Dans le cadre d'OGIIS, il peut-être rappelé qu'une réticence a priori de la part de certains contributeurs potentiels à prendre en main des outils ne relevant pas de leur spécialité et la crainte des « usines à gaz » sont des faits observés depuis le début des concertations relatives aux bases de données des OHM. La réponse à ces craintes se situe certainement dans l'évolutivité et la convivialité des outils à concevoir.

Cela appelle des méthodes de conception qui pour être adéquate devront solliciter différents types d'utilisateurs à divers degrés d'implication : le pilotage d'un tel outil par et pour les OHM devra être à la fois consultatif, participatif et éclairé. Autrement dit, les échanges humains seront les meilleurs garants de la performance informatique.

ACRONYMES / SIGLES

APR : Appel à Proposition de Recherche

BBEES : Bases de données sur la Biodiversité, Ecologie, Environnement et Sociétés

CNRS : Centre National de la Recherche Scientifique

DRIIHM : Dispositif de Recherche Interdisciplinaire des Interactions Hommes-Milieus

GN : GeoNetwork

Huma-Num : Humanités Numériques

INEE : Institut Ecologie et Environnement

INSHS : Institut des Sciences Humaines et Sociales

INSPIRE : Infrastructure for Spatial InfoRmation in Europe

ISO : International Organization for Standardization

LabEx : Laboratoire d'Excellence

MRCT : Mission Ressources et Compétences Technologiques

MI : Mission Interdisciplinarité

OGIIS : Outil de Gestion Intégré de l'Information à caractère Scientifique

OHM : Observatoire Hommes-Milieus :

- **BMP** : Bassin Minier de Provence
- **EST** : Estarreja
- **GMV** : Tessékéré
- **HV** : Haut-Videssos
- **LM** : Littoral méditerranéen
- **NUN** : Nunavik
- **OYA** : Oyapock
- **VR** : Vallée du Rhône

OHMI : OHM Internationaux

OS : Système d'Exploitation (Operating System en anglais)

RBDD : Réseau Bases De Données

ROHM : Réseau des OHM

SI : Système d'Information

TGIR : Très Grande Infrastructure

ZA : Zone Atelier

ZABR : Zone Atelier du Bassin du Rhône

BIBLIOGRAPHIE

- ASSEMBLEE NATIONALE (France), 1997. Loi n° 97-283 portant transposition dans le code de la propriété intellectuelle des directives du Conseil des Communautés européennes n° s 93/83 du 27 septembre 1993 et 93/98 du 29 octobre 1993 - TITRE II - [art. L123-1 \(V\)](#) at <http://www.legifrance.gouv.fr/affichCodeArticle.do;jsessionid=FAAC5EF1CDC709D72A236B8B9DE8C0C7.tpdjo17v_2?cidTexte=LEGITEXT000006069414&idArticle=LEGIARTI000006278937&dateTexte=20130923&categorieLien=id#LEGIARTI000006278937>
- COMMISSION EUROPÉENNE, 2012. Proposition de RÈGLEMENT DU PARLEMENT EUROPÉEN ET DU CONSEIL relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données (règlement général sur la protection des données).
- DUJARDIN Florence, 2012. Le projet de gestion l'information de l'EMSE. Journée OGIIS « Métadonnées et information scientifique », Aix-en-Provence
- FOLIOT Gérald, 2012. ISIDORE, un outil qui collecte, enrichit et offre un accès unifié aux données numériques des SHS. Journée OGIIS « Métadonnées et information scientifique », Aix-en-Provence
- GBIF, 2012. Comment le DwC a changé la manière de publier les données de biodiversité présenté par Michael Akbaraly lors de la Formation sur la publication des données de biodiversité dans le réseau GBIF et leur aptitude à être utilisées, Paris. URL : <http://www.gbif.fr/wp-content/uploads/2012/11/02-2011-09-28-How-the-Darwin-Core-Archives-have-changed-the-landscape-of-biodiversity-data-publishing1.pdf>, dernière consultation le 23/09/2013
- HEINTZ Wilfried, 2012. Métadonnées & catalogage. Retour d'expérience - projet CaSIF. Journée OGIIS « Métadonnées et information scientifique », Aix-en-Provence
- INEE, 2013. Prospective de l'Institut Ecologie et Environnement du CNRS, [Compte-rendu des journées des 24 et 25 octobre 2012](#), Avignon, Les Cahiers Prospectives (hors-série), 244 pages.
- INEE, 2009. Compte-rendu des journées de prospective de l'Institut Ecologie et Environnement (INEE) du CNRS, Rennes, 35 pages.
- LABORIE P., 2006. Cycle de vie d'un projet scientifique, passage de la R&D à la réalisation.
- LAGADEUC Yvan, CHENORKIAN Robert, 2009. Les systèmes socio-écologiques : vers une approche spatiale et temporelle. *Natures Sciences Sociétés* **17**, pages 194-196. DOI: 10.1051/nss/2009032
- MERRIEN Francis, LEOBET Marc, 2011. La directive Inspire pour les néophytes, 63 pages.
- MICHENER W. K., JONES M. B., 2012. Ecoinformatics: supporting ecology as a data-intensive science. *Trends in Ecology & Evolution* **27**, pages 85–93.
- MOREL-PAIR Catherine, 2012. Métadonnées pour les bibliothèques numériques. Enjeux, standards et pratiques. INIST-CNRS, 95 pages.
- PARLEMENT EUROPÉEN & CONSEIL DE L'UNION EUROPÉENNE, 2007. DIRECTIVE du 14 mars 2007 établissant une infrastructure d'information géographique dans la Communauté européenne

(INSPIRE). 2007/2/CE 14 (2007). at <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2007:108:0001:0014:FR:PDF>

OSU de Rennes, 2011. Séminaire Métadonnées : Concepts et Outils. Université de Rennes

TASSIN Philippe (dir.), 2010. Systèmes d'information et développement durable. Economie, société et environnement. Collection Études informatiques et logiciels. Lavoisier, 345 pages.

VAN NOORDEN Richard, 2013. Everything on display. *Nature* **500**, pages 243-245.

|

ANNEXES

LISTE INDICATIVE ET NON-EXHAUSTIVE DES CORRESPONDANCES ENTRE ELEMENTS DE 2 MODELES DE DESCRIPTION DE RESSOURCES : DC ET INSPIRE

Dublin Core	INSPIRE	Description commune
[en] [fr] [code]	[en] [fr] [code (classe)]	[Explications (convention de saisie)]
Identifiant Identifiant dc:identifiant	File identifier Désignation de la donnée (URI) gmd:dataSetURI	Chaîne de caractères unique identifiant la ressource sans ambiguïtés (exemple : PI_SIRET_AUTEURAAAA_rowid).
Language Langue dc:language	Language Langue gmd:language	Langue(s) employée(s) dans la ressource et sa description (exprimé par un code issu de la liste ISO639-2 bibliographique).
Creator Créateur dc:creator	Originator A l'origine de gmd:CI_RoleCode codeListValue='originator' (CI_ResponsibleParty)	Personne responsable de la création de la ressource.
Publisher Editeur dc:publisher	Distributor Distributeur gmd:distributorContact (CI_ResponsibleParty, MD_Distributor)	Personne responsable de la diffusion de la ressource.
Contributor Contributeur dc:contributor	Processor Personnes responsables du processus gmd:processor (CI_ResponsibleParty)	Personne impliquée dans la production ou la diffusion de la ressource.
Title Titre dc:title	Title Titre gmd:title (CI_Citation)	Titre sous lequel la ressource est connue
Date Date dc:date	Date Date gmd:date (CI_Citation, CI_Date)	Date de référence pour la ressource en question

Description Description dc:description	Abstract Résumé gmd:abstract	Description du contenu de la ressource sous forme de résumé ou de liste structurée
Subject Sujet dc:subject	Keyword Mot clé gmd:keyword	Mots-clés ou thèmes, de préférence issus de vocabulaire contrôlés (thésaurus) relatifs au sujet de la ressource
Coverage Couverture dc:coverage	Extent Etendue gmd:extent (Ex_Extent, DQ_Scope)	Informations relatives à l'extension spatiale et temporelle des données
Provenance Provenance dc:provenance	Lineage Généalogie gmd:statement (LI_Lineage, DQ_Dataquality)	Explication générale de la filiation du jeu de données.
Type Type dc:type	Presentation Form Forme de la présentation gmd:presentationForm	Nature ou du genre de contenu de la ressource selon le DCMI Type Vocabulary .
Rights Droits dc:rights	Resource constraints Contraintes sur la ressource gmd:resourceConstraints (MD_constraints)	Contraintes en matière d'accès et d'utilisation.
Format Format dc:format	Resource format Format de la ressource gmd:resourceFormat (MD_Format)	Description du format de la ressource selon un vocabulaire contrôlé (voir http://www.isi.edu/[...]/iana/assignments/media-types/media-types).
Relation Relation dc:relation	Other citation details Autres informations de référence gmd:otherCitationDetails	Autre ressource liée, utilisée pour compléter les informations de référence sous la forme d'une chaîne de caractère relevant d'un système d'identification formel (URL, URI, côte d'archive, etc...).
Source Source	Source Source	Ressource à partir de laquelle la ressource décrite est dérivée, référencées sous la forme d'une adresse ou d'un identifiant

dc:source	gmd:source (LI_Source)	uniforme (URL ou URI).
Audience Audience dc:audience	Distribution information Informations sur la distribution gmd:distributionInfo	Indications concernant le contexte de réutilisation des données.

Sources :

- <http://dublincore.org/documents/dces/>
- <http://www.bibl.ulaval.ca/dublincore/usageguide-20000716fr.htm>
- http://www.bnf.fr/fr/professionnels/formats_catalogage/a.f_dublin_core.html
- <http://geonetwork-opensource.org/manuals/trunk/fra/users/appendix/glossary/iso19139/index.html>

FORMULAIRE D'INSCRIPTION A OGIIS (VUE FACTICE)



[CNRS](#) | [INEE](#) | [Annuaire](#) | [Mots-Clefs CNRS](#) | [Autres sites](#)

Observatoires Hommes-Milieu de l'INEE
Centre national de la recherche scientifique



Inscription

[obligatoire]

Nom :

[obligatoire]

Prénom :

[obligatoire]

Organisme :

Unité :

Quartier :

Numéro de voie :

Type de voie :

Nom de voie :

Bâtiment :

Boîte postale :

Code postal :

[obligatoire]

Ville :

[obligatoire]

Cedex :

Pays :

Téléphone :

[obligatoire]

e-mail :

Spécialités :

Bref CV :

[obligatoire]

Login :

[obligatoire]

Mot de passe :

[obligatoire]

Répétez le mot de passe :

Saisissez le résultats de l'opération.

[obligatoire]

4 + 1 =

FORMULAIRE DE REPONSE AUX APR (VUE FACTICE)



[suivant](#) ►

OHM concerné(s) *[Obligatoire]*

Veillez cocher le ou les OHM concerné(s) par votre proposition.

- réseau des OHM
- OHM Bassin minier de Provence
- OHM Oyapock
- OHM Haut-Videssos
- OHM Tessékéré
- OHM Estarreja
- OHM Vallée du Rhône
- OHM Littoral méditerranéen

1. TITRE du projet

Titre complet *[Obligatoire]*

Sigle



[◀ précédent](#) [suivant ▶](#)

2. Responsable(s)

2.1 Description détaillée du porteur principal

<i>[obligatoire]</i>	Nom :	<input type="text"/>
<i>[obligatoire]</i>	Prénom :	<input type="text"/>
<i>[obligatoire]</i>	Organisme :	<input type="text"/>
	Unité :	<input type="text"/>
	Quartier :	<input type="text"/>
	Numéro de voie :	<input type="text"/>
<i>[obligatoire]</i>	Type de voie :	<input type="text"/>
	Nom de voie :	<input type="text"/>
	Bâtiment :	<input type="text"/>
	Boîte postale :	<input type="text"/>
<i>[obligatoire]</i>	Code postal :	<input type="text"/>
<i>[obligatoire]</i>	Ville :	<input type="text"/>
	Cedex :	<input type="text"/>
<i>[obligatoire]</i>	Pays :	<input type="text"/>
<i>[obligatoire]</i>	Téléphone :	<input type="text"/>
<i>[obligatoire]</i>	e-mail :	<input type="text"/>

[obligatoire]

Bref CV

B *I* « » “ ”



◀ [précédent](#) [suivant](#) ▶

2.2 Description détaillée du co-porteur

Nom :	
Prénom :	
Organisme :	
Unité :	
Quartier :	
Numéro de voie :	
Type de voie :	
Nom de voie :	
Bâtiment :	
Boîte postale :	
Code postal :	
Ville :	
Cedex :	
Pays :	
Téléphone :	
e-mail :	

Bref CV

B *I*   « » “ ”



◀ [précédent](#) [suivant](#) ▶

3. Participants (si le projet est présenté par un collectif)

Nom	Prénom	Fonction	Organisme	Unité	Téléphone	Courriel



◀ précédent suivant ▶

4. Exposé de la proposition

Domaine(s) disciplinaire(s) concerné(s) *[Obligatoire]*

Cadrage problématique *[Obligatoire]*

B *I*   « »

Mot(s) clé(s) *[Obligatoire]*

Séparer par des points virgules ";"

Objectifs *[Obligatoire]*

B *I*   « »

Résultats escomptés *[Obligatoire]*

B *I*   « »

Méthodologie proposée *[Obligatoire]*

B *I*   « »

Motivation du traitement de la proposition dans le cadre du ROHM, du et/ ou des OHM concernés.

[Obligatoire]

B *I*   « »



[← précédent](#)

6. Validation

J'affirme avoir lu le texte de l'APR auquel répond la proposition ci-dessus et m'engage en cas d'acceptation de mon projet de recherche à signer la charte du ou des OHM concerné(s). *[Obligatoire]*

- Oui
 Non

FORMULAIRE DE SELECTION DES REPONSES AUX APR (VUE FACTICE)



4 propositions de recherche adressées à votre OHM (APR 2014)

Porteur principal	Titre	Statut
NOM Prénom	Intitulé complet (SIGLE)	retenu
OPN Qsfopn	Joujuvmf dponqmfu (TJHME)	refusé
PQO Rtgpqo	Hpvkvwng eqorngv (UKNG)	<input type="text"/>
QRP Suhqrp	lqwjwxoh fhpsohw (VLJOH)	clos



Valider

FORMULAIRE DE SAISIE DES RAPPORTS DE RESTITUTION (VUE FACTICE)



[suivant](#) ▶

Rechercher une proposition de recherche

Indiquez le titre, l'année ou l'auteur du document

Choisissez, parmi les résultats, l'étude à restituer.

Liste des résultats avec lien vers la proposition complète et bouton radio pour sélectionner l'étude



[suivant](#) ▶

Votre projet

OHM concerné(s)

- réseau des OHM
- OHM Bassin minier de Provence

1. TITRE du projet

Titre complet

Territory, an Influential Factor in understanding the level of Environmental Awareness of children

Sigle

Eaif

2. Responsable(s)

2.1 Description détaillée du porteur principal

NOM Prénom, fonction, EPST – UMR 0000 SIGLEUMR, 0 voie Nom de voie 13000
MARSEILLE, tél. : +33 400 000 000, mèl : personnage@domaine.fr



2.2 Description détaillée du co-porteur

NOM Prénom, fonction, EPST – UMR 0000 SIGLEUMR, 0 voie Nom de voie 13000
MARSEILLE, tél. : +33 400 000 000, mèl : personnage@domaine.fr



3. Participants (si le projet est présenté par un collectif)

Nom	Prénom	Fonction	Organisme	Unité	Tél.	Courriel
Ogn	Qestgn	Chercheur	EPST	SOR 0000 SSOR	+330000000	np@o.fr



4. Bilan scientifique

4.1 Actions engagées (5000 caractères maximum)

B *I*     « »

4.2 Description du procédé d'observation

B *I*     « »

4.3 Résultats scientifiques (10 000 caractères maximum)

B *I*     « »

4.4 Retombées (avérées ou projetées)

B *I*     « »

4.5 Remarques diverses

B *I*     « »



dépasser les frontières



◀ précédent suivant ▶

5. Financement

Budget total du projet proposé *[Obligatoire]*

Le projet fait-il l'objet de demande de financements complémentaires ?

- Oui
- Non

S'inscrit-il dans un programme de recherche plus vaste ?

- Oui
- Non

Si oui, précisez ces financements et le programme.



Montant accordé

Détail des dépenses

Intitulé	Montant	Nature	
		<input type="text"/>	+ X
		<input type="text"/>	+ X
		<input type="text"/>	+ X



[← précédent](#)

7. Validation

Aperçu des informations saisies

Valider

MODELE LOGIQUE DE DONNEES

