



HAL
open science

Construire des outils pour la gestion des données de la recherche dans une communauté d'universités

Aurore Cartier, Magalie Moysan, Nathalie Reymonet

► To cite this version:

Aurore Cartier, Magalie Moysan, Nathalie Reymonet. Construire des outils pour la gestion des données de la recherche dans une communauté d'universités. Journée sur les données de la recherche, ADBS, Jan 2015, Paris, France. hal-01138663

HAL Id: hal-01138663

<https://hal.science/hal-01138663v1>

Submitted on 7 Apr 2015

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.



Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

Construire des outils pour la gestion des données de la recherche dans une communauté d'universités

A. Cartier, M. Moysan, N. Reymonet

Journée sur les données de la recherche, ADBS, 16/01/2015

Aurore CARTIER, conservateur au Service commun de la documentation de l'université Paris Descartes

Magalie MOYSAN, archiviste au Bureau des archives de l'université Paris Diderot

Nathalie REYMONET, documentaliste à la Direction d'appui à la recherche et à l'innovation de l'université Paris Diderot

INTRODUCTION

La gestion des données représente un enjeu économique et sociétal important, comme en témoigne l'émergence d'expressions telles que *Open Data* et *Big Data*. Dans le domaine scientifique, le partage et l'exploitation des données est un remarquable gisement pour de nouvelles recherches, et le lancement du projet pilote « *Open data research pilot* » du programme Horizon 2020 ouvre de nouvelles perspectives de travail aux professionnels de l'information scientifique et technique (IST).

Dans ce contexte, comment ces derniers peuvent-ils apporter leurs compétences aux chercheurs dans la gestion, la description, la valorisation et la conservation de l'information numérique ?

Notre groupe de travail, qui réunit de façon informelle des compétences en matière de gestion de l'information, de documentation et d'archivage, est issu des universités Paris Descartes et Paris Diderot. Cette complémentarité des compétences s'est vite révélée être un atout, tant les problématiques de partage, de dissémination et de conservation des données se complètent. La question du positionnement de ce groupe s'est rapidement posée : quelle place occuper en tant que professionnels de l'IST, dans l'accompagnement des projets de recherche à l'échelle d'une ComUE (Communauté d'université et d'établissement) naissante ? A notre sens, notre expertise se situe dans l'accompagnement méthodologique vers davantage de rigueur et de formalisme dans la gestion des données *de* et *sur* la recherche.

Le lancement du projet pilote de la Commission européenne (CE) a été un élément déclencheur de la démarche adoptée par notre ComUE Sorbonne Paris Cité, qui réunit quatre universités -Sorbonne Nouvelle, Paris Descartes, Paris Diderot et Paris 13- et quatre établissements de recherche -Sciences Po, Institut de Physique du Globe de Paris, INALCO, Ecole de Hautes Etudes en Santé Publique, ainsi que des organismes de recherche nationaux. La ComUE rassemble en effet les ingénieurs-projets Europe de ses membres en un réseau de compétences, la Cellule Europe SPC.

Notre démarche a donc été de fournir aux chercheurs et ingénieurs-projets de nos établissements des outils pratiques et harmonisés, qui soient susceptibles de répondre aux exigences des organismes financeurs, au-delà de la seule CE, et de constituer une base de coopération avec les organismes de recherche partenaires. Dans ce cadre, la gestion des données se traduira par une description, un traitement, un entreposage, voire une sélection des données, et les outils fournis doivent guider les chercheurs dans la réalisation de ces actes.

En coopération avec le responsable information scientifique et technique (IST) et la cellule Europe de SPC, nous avons donc commencé à monter :

- un modèle de *Data Management Plan* (DMP), ou plan de gestion des données, accompagné d'une note explicative sur les enjeux liés aux données de la recherche, modèle publié depuis le 9 janvier dernier ; le DMP est un livrable pour la Commission européenne (CE) ;
- un modèle de plan de dissémination des résultats de la recherche, qui est également un livrable pour la CE ;
- un plan de formation à destination des ingénieurs-projets Europe et des chercheurs au sein de SPC, sur la gestion et la dissémination des données de la recherche et sur la rédaction de DMP dans le cadre d'Horizon 2020.

L'outil d'aide à la rédaction de *Data Management Plan*, le modèle de plan de dissémination et le plan de formation sont un point d'entrée dans le projet de recherche pour les professionnels de l'IST que nous sommes. C'est également un moyen de confronter notre expertise théorique à la pratique en intervenant dès la phase de montage des projets de recherche, en soutien aux ingénieurs-projets. Ceux-ci constituent un élément essentiel de la stratégie adoptée, en tant que référents des chercheurs dans les appels à projets. Enfin, la construction atypique de notre groupe de travail et nos rattachements hiérarchiques variés (SCD, Direction de la recherche et Secrétariat général), représentent un atout pour l'accompagnement des chercheurs dans la gestion de leurs données, en offrant l'opportunité de relayer tant les besoins que les réponses apportées à différents niveaux et auprès de tous les acteurs de la recherche.

1. CIRCONSTANCES ET OPPORTUNITES

Le programme cadre de la CE, Horizon 2020, a été lancé en décembre 2013 et sera clôt en 2018. Il présente plusieurs nouveautés, dont l'affirmation de la position de la CE sur l'*open access* (libre accès) des publications scientifiques et l'ouverture des données de la recherche.

L'Open Access dans Horizon 2020

Le schéma « *OA in context : dissemination & exploitation of research results* » présente la position de la Commission européenne sur les résultats de la recherche qu'elle finance dans Horizon 2020.

La recherche donne lieu à des résultats, qui vont être accompagnés de deux informations : d'une part le plan de dissémination, qui décrit ce que le coordinateur du projet prévoit de faire en termes de diffusion des résultats, et d'autre part le plan de gestion des données, qui établit les responsabilités et les mesures qui seront prises pour la gestion des données issues du projet.

Deux voies principales sont possibles : d'une part, la diffusion et le partage et d'autre part, l'exploitation économique et la protection, par exemple par dépôt de brevet.

Si les résultats sont diffusés sous forme de publications, celles-ci seront mises en ligne en *open access* en conformité avec l'article 29.2 du *Grant Agreement* -la convention de financement entre la Commission européenne et l'établissement de l'équipe lauréate.

Les équipes de recherche peuvent publier dans une revue -ce qu'on appelle *gold open access*-, puisque les *Articles Processing Charges* -frais de publication- sont éligibles pendant la durée du projet, ou rendre publics leurs travaux dans une archive ouverte -que l'on appelle *green open access*-

avec le dépôt des métadonnées associées aux publications –les métadonnées décrivent les documents déposés. Ce dépôt en archive ouverte doit éventuellement respecter une période d’embargo -délai pendant lequel une information est retenue.

Si les résultats de la recherche se traduisent aussi par la production de données, il y aura dépôt de ces données, en conformité avec l’article 29.3 du *Grant Agreement* ; ces données peuvent être déposées et rendues accessibles librement, ou protégées pour toutes sortes de raisons. Dans tous les cas, elles seront décrites dans un *Data Management Plan* (DMP).

Groupe de travail Paris Descartes / Paris Diderot

Le Service commun de la documentation de l’université Paris Descartes, le Bureau des archives et la Direction d’appui à la recherche et à l’innovation de l’université Paris Diderot, ont proposé des formations et des outils à la Cellule Europe SPC qui regroupe les ingénieurs-projets Europe de la ComUE Sorbonne Paris Cité. Ces outils consistent en un modèle de DMP et un modèle de plan de dissémination des résultats d’un projet européen.

L’Open Research Data Pilot

Dans le cadre de Horizon 2020, la Commission Européenne a lancé fin 2013 un projet pilote qui porte sur les données de la recherche, l’*Open Research Data Pilot*. Les lauréats des appels à projet figurant dans ce pilote doivent :

- déposer les données issues du projet dans un entrepôt (conformément à l’article 29.3 du *Grant Agreement*). Ce pilote concerne les données (et métadonnées associées), nécessaires à la validation des résultats présentés dans les publications, ainsi que les autres données (et métadonnées associées) décrites dans le plan de gestion de données ;
- fournir des informations nécessaires à la validation des résultats, documenter ces données et en particulier prendre les mesures nécessaires pour rendre ces données accessibles et exploitables librement, par le biais de licences telles que *Creative Commons*. Par exemple, il pourra s’agir de préciser les outils et le matériel nécessaires à la validation des résultats publiés, comme des logiciels, des algorithmes, ou des protocoles d’analyse. Dans la mesure du possible, les bénéficiaires doivent aussi fournir les outils qui ont servi à produire ces données
- produire un plan de gestion de données qui a pour vocation d’anticiper les modalités de gestion et de dissémination des données. La première version sera envoyée à la Commission européenne dans les six mois qui suivent le lancement du projet. Le DMP sera mis à jour au minimum à mi-parcours du projet et lors de l’évaluation finale du projet, au moment de sa clôture.

Le périmètre de l’Open Research Data Pilot

Nous reprenons la représentation Horizon 2020 en trois piliers qui correspondent aux trois priorités définies par la Commission européenne :

- renforcer l’ « Excellence scientifique », pilier portant sur la recherche fondamentale,

- développer un leadership industriel : « Primauté industrielle », pilier portant sur la recherche finalisée,
- relever les « Défis sociétaux », pilier portant sur la recherche ciblée.

Sur cette représentation, nous avons surligné en orange, la couleur de l'open access, les thématiques couvertes par le pilote sur les données. Ce pilote vise ainsi les thématiques portant sur les nouvelles technologies ou qui génèrent d'importantes masses de données, ainsi que les questions transverses de société.

Il est à noter que les appels de l'ERC (*European Research Council*) ne font pas partie du pilote sur les données, mais que la CE a émis des recommandations: « ... l'ERC soutient le principe d'ouverture de données, et recommande à ses chercheurs de suivre les bonnes pratiques en conservant les fichiers de toutes les données de recherche qu'ils ont utilisées pendant leurs travaux, et qu'ils se préparent à partager leurs données avec d'autres chercheurs... » .

Les solutions d'appui à H2020

Pour la dissémination des résultats, la Commission met à la disposition des chercheurs l'entrepôt de données scientifiques ZENODO, issu du projet européen OpenAIRE+ et développé par le CERN. S'il n'existe pas d'entrepôt thématique approprié, ZENODO peut être utilisé pour le dépôt des données scientifiques. Cet entrepôt affecte des identifiants pérennes aux objets et expose les métadonnées de description au moissonnage à travers le protocole OAI-PMH, le protocole d'interopérabilité des archives ouvertes, utilisé par exemple par l'archive nationale française, HAL.

Pour la préservation, le projet EUDAT est une initiative européenne cofinancée par le 7e PCRD. Il vise à répondre aux besoins futurs des chercheurs en matière d'accès et de préservation des données scientifiques, et propose des services tels que la réplication sécurisée des données, le transfert des données vers des centres de traitement, le service d'archivage de petits fichiers, un catalogue de métadonnées, une infrastructure d'authentification et d'autorisation. Le CINES participe à ce projet européen comme centre de ressources.

2. L'ELABORATION D'UN PLAN DE GESTION DE DONNEES

Pourquoi gérer les données ?

Aujourd'hui, que deviennent les données d'un projet de recherche ? On peut représenter les données d'un projet sous la forme d'un entonnoir :

- le chercheur produit ou collecte un grand nombre de données primaires,
- il n'en traitera qu'une partie car tout ne lui sera finalement pas utile,
- et seule une petite partie des données seront exploitées pour donner lieu à une publication.

Actuellement, les données de recherche accessibles se réduisent aux données intégrées dans les publications : schémas, tableaux, etc... Seule une petite partie des données de recherche est visible, le reste est invisible à la communauté scientifique. Dans la mesure où elles ne sont pas diffusées, ni archivées, ces données terminent bien souvent leur vie sur un disque dur et deviennent illisibles, et sont donc perdues.

Il ne s'agit évidemment pas de tout conserver, mais de sélectionner les données à conserver et diffuser.

Plusieurs constats

Les données ne font pas l'objet d'une valorisation propre. En effet, rendre les données accessibles ne signifie pas qu'elles soient réutilisables, il existe plusieurs types d'obstacles :

- obstacles techniques : les données sont souvent intégrées dans les publications sous forme de graphiques ou d'illustrations (publications statiques), et rarement publiées sous forme de données brutes. Les données ainsi sont rarement identifiées comme des éléments distincts du reste de la publication ;
- obstacles juridiques : si un document n'est pas sous une licence autorisant sa réutilisation, même son auteur peut ne pas avoir le droit de le réutiliser ; de plus, le statut juridique des données est souvent flou aux yeux des producteurs eux-mêmes ainsi que des utilisateurs ;
- obstacles intellectuels : il manque souvent le contexte du cadre de production des données, car celui-ci n'est pas toujours suffisamment décrit dans la publication.

En conséquence, une part importante des données produites par la recherche reste inutilisable.

Par ailleurs, le sort des données après leur phase d'utilisation courante est souvent négligé : tri, destruction, archivage intermédiaire, archive définitif etc. alors même qu'une bonne gestion des données peut devenir un atout pour la recherche, actuelle et future.

Le *Data Management Plan* : anticiper le sort des données

Dans le cadre d'H2020, la Commission européenne donne la définition suivante du DMP : *"A DMP describes the data management life cycle for all data sets that will be collected, processed or generated by the research project. It is a document outlining how research data will be handled during the research project, and even after the project is completed, describing what data will be collected, processed or generated and following what methodology and standards, whether and how data will be shared and/or made open, and how it will be curated and preserved."*

Ainsi, le DMP décrit le cycle de gestion de toutes les données qui seront collectées, traitées ou générées par un projet de recherche. Il a pour vocation d'anticiper les problèmes de gestion qui peuvent survenir au cours d'une recherche et les conditions d'une conservation permettant une diffusion future des données.

Les étapes de rédaction et de validation du DMP dans Horizon 2020

Le DMP est un document qui accompagne la gestion des données d'un projet de recherche. Il s'inscrit donc dans la temporalité du projet de recherche et est amené à évoluer avec lui. Dans le cadre du projet pilote de H2020, le DMP n'est pas requis pour répondre à l'appel à projets mais seulement si le projet est accepté.

Dans la réponse à l'appel à projets, le chercheur doit renseigner la Commission européenne sur le fait qu'il va ou non réaliser un plan de gestion de données. Il peut entrer dans le pilote s'il n'en fait pas partie (*opt-in*) ou en sortir s'il en fait partie, mais en le justifiant (*opt-out*).

Une fois son projet accepté, le chercheur prépare la rédaction du DMP en collaboration avec l'ingénieur-projet, puisqu'il doit proposer une première version dans les 6 premiers mois du projet. Pour l'aider dans cette tâche, et selon un workflow que nous proposons également, il peut être aidé par différents professionnels dans cette rédaction : bibliothécaires, documentalistes, archivistes, informaticiens.

L'étape de mise en place des outils de gestion consiste en particulier en l'établissement d'un plan de classement, la définition de règles de nommage et de procédures.

Le DMP peut être révisé au cours du projet. La CE recommande de faire au moins trois versions du DMP : dans les 6 premiers mois, à *mid-term* et au moment du rapport final. Sur un projet complexe, il y aura potentiellement de plus nombreuses versions, ainsi que de nombreux jeux de données décrits.

A la fin du projet, le chercheur doit rendre la version finale de son DMP, en même temps que tous les documents finaux demandés par la Commission, comme par exemple la liste des publications.

Guide de rédaction de plans de gestion de données

Nous avons conçu ce document afin d'accompagner les chercheurs et ingénieurs-projets lors de la rédaction du DMP : sa structure s'appuie sur le modèle proposé par la Commission européenne dans le cadre d'Horizon 2020 mais aussi sur plusieurs modèles existants tels que celui de la *National Science Foundation* (NSF) ou de l'*Interuniversity Consortium for Political and Social Research* (ICPSR).

Le workflow des acteurs (description du flux d'information entre les acteurs) est détaillé après les différentes sections du DMP. Celui-ci est structuré en cinq sections.

Sections 1 à 3 : Informations sur le projet et la responsabilité du plan de gestion

Ces trois sections du DMP se situent à l'échelle du projet (et non des jeux de données). Un jeu de données peut désigner : i) un lot techniquement homogène (formats informatiques des données homogènes entre-eux), ou ii) un lot intellectuellement cohérent même si celui-ci est composé de lots techniquement hétérogènes.

Les objectifs sont d'informer sur le contexte du projet de recherche et de définir le suivi des données. Les principaux champs de ces sections sont :

- Identifiant de l'appel à projets, thématiques et autres éléments d'ordre administratif concernant le *projet* de recherche
- Informations concernant la responsabilité et la propriété des données issues de ce projet
- Ressources nécessaires à la mise en œuvre du DMP, en termes de matériel, de personnels et de formations éventuellement nécessaires, et les montants financiers que cela représente.

Section 4.1 : Description des données

La section 4 se situe à l'échelle du jeu de données (*dataset*) et peut donc être dupliquée pour chaque jeu de données, autant de fois que nécessaire à la description de tous les jeux de données. Les objectifs sont de présenter le type de données qui seront produites et reçues dans le cadre du projet. Les principaux champs sont :

- Identifiant du jeu de données

- Nature des données
- Réutilisation de données existantes, le cas échéant
- Méthode de production des données
- Formats des données.

Section 4.2 : Stockage, accès et sécurité des données

Les objectifs sont de définir les modalités d'hébergement, de sauvegarde et d'accès aux données pendant la phase active du projet. Les principaux champs sont :

- Volumétrie prévisionnelle de données produites par le projet
- Modalités de gestion des accès à ces données
- Identification des risques et menaces telles que confidentialité, intégrité, traçabilité, disponibilité des données
- Prévention des risques ainsi identifiés.

Section 4.3 : Métadonnées, documentation et organisation des données

Les objectifs sont de préciser la manière dont seront décrites et organisées les données produites ou reçues au cours du projet. Les principaux champs sont :

- Standards et formats disciplinaires existants, le cas échéant
- Mode de production et responsabilité des métadonnées
- Règle de nommage adoptée pour le jeu de données
- Arborescence de classement choisie
- Autres informations complémentaires nécessaires à la documentation et à l'organisation des données.

Section 4.4 : Dissémination des données

Les objectifs sont de préciser les modalités et les éventuelles précautions éthiques, juridiques et techniques selon lesquelles les données seront diffusées. Les principaux champs sont :

- Principe général de diffusion des données auquel répond le DMP
- Justification de la protection des données sensibles, le cas échéant
- Licence utilisée (*Creative Commons*, GNU...)
- Nom de l'entrepôt choisi pour le dépôt et la mise à disposition éventuelle des données
- Existence de publications associées
- Potentiel de réutilisation des données
- Durée éventuelle de l'embargo sur les données décrites.

Section 5 : Sélection et archivage

Les objectifs sont de renseigner sur le sort des données à l'issue du projet et les dispositifs de conservation choisis. Les principaux champs sont :

- Mode de sélection des données
- Sort des données à l'issue du projet

- Volume final des données conservées
- Durée de conservation des données sélectionnées pour la conservation
- Nom de la plateforme d'archivage choisie pour les données sélectionnées.

Cette section concerne l'ensemble des données produites ou collectées au cours du projet, qu'elles aient été diffusées ou non.

3. COMPETENCES ET ACTEURS DANS LA MISE EN ŒUVRE D'UN DMP

Les acteurs ressources du plan de gestion de données

Nous avons conçu une présentation de la mise en place du DMP comme un workflow entre les différents acteurs, de façon à décrire les acteurs du plan de gestion de données et la chronologie dans laquelle ils interviennent. La constitution d'un DMP fait en effet appel à des compétences multiples : chercheur, ingénieur-projets, informaticiens, professionnels de l'information scientifique et technique. Ces derniers peuvent être selon les cas et les établissements, des documentalistes, des bibliothécaires, ou des archivistes.

Chercheur

Le DMP concerne l'ensemble du projet financé, et sa responsabilité incombe à celui qui sera désigné par le coordinateur du *projet* comme le responsable des *données*. Ce responsable des données, ou coordinateur du DMP, est le chercheur en charge de la gestion des données : de notre point de vue, il doit être choisi pour avoir une vue suffisamment transversale dans le projet à la fois sur tous les types de données et sur toute la durée du projet : il est responsable de la création du DMP et de sa mise en œuvre au cours du projet. Ce n'est clairement pas le rôle d'un stagiaire.

Ingénieur projet

L'ingénieur projet, ou chargé de projet, va intervenir aux côtés du chercheur dès le début puisqu'une partie du contenu du DMP existe déjà, avec des informations produites par différents acteurs à différents moments, et seront regroupées dans le document unique qu'est le DMP : un certain nombre d'informations sur le projet, qui concernent la responsabilité des données et les ressources nécessaires à la mise en place du DMP ont déjà été prévues lorsque le chercheur a préparé la partie budgétaire de son projet pour répondre à l'appel à projet ou se trouvent dans les différents documents d'accompagnement tels que l'accord de consortium. L'ingénieur projet est un acteur important qui peut aider le chercheur dans cette démarche descriptive en ce qu'il peut accéder à un grand nombre de ces informations. Pour un établissement coordinateur du projet, l'ingénieur-projets sera la personne clé pour la coordination des actions autour du DMP. Par exemple, parmi les coûts à prendre en compte dans la mise en œuvre du DMP, il peut y avoir des coûts de formation pour acquérir les compétences nécessaires au traitement des données. Ces coûts sont éligibles pour la CE, il est important que l'ingénieur-projets en tienne compte lors de l'établissement du budget du projet. Il y a donc une interaction assez étroite entre le chercheur responsable du DMP et l'ingénieur-projets pour les trois premières parties du DMP.

Description des données

Le chercheur sait ce qu'il va produire, il est donc le seul à pouvoir décrire les données produites et l'éventuel découpage entre jeux de données. Un jeu de données peut désigner : i) un lot

techniquement homogène, ou ii) un lot intellectuellement cohérent même si celui-ci est composé de lots techniquement hétérogènes.

Stockage accès, sécurité des données

Des informaticiens seront sollicités pour répondre ces questions. (informaticien de laboratoire ou DISI/DSI, cela dépendra des structures au cas par cas).

Métadonnées

Les métadonnées sont un ensemble de données structurées décrivant des ressources physiques ou numériques. Elles sont classiquement divisées en métadonnées descriptives, métadonnées administratives et métadonnées de structure. Par exemple pour une publication : le document lui-même sera « les données », et la référence bibliographique sera « les métadonnées ». Dans ce cas, les auteurs, la date de publication ou le nombre de pages sont des éléments administratifs, alors que le titre et les mots-clés sont des éléments descriptifs du contenu. Pour les métadonnées, il existe déjà des standards dans certaines disciplines. A défaut, les référents IST peuvent venir en appui des chercheurs pour leur proposer des tables de correspondance (*mapping*) de métadonnées en fonction des besoins liés à la dissémination ou à l'archivage des données, si le chercheur sait déjà dans quel entrepôt il va déposer les données.

Dissémination

Pour les aspects de diffusion des données, on va pouvoir reprendre des éléments déjà mentionnés dans la réponse à l'appel à projets : le chercheur sait dans quel journal il va vouloir publier, en revanche, c'est plutôt à l'IST de vérifier le type d'open access possible pour telle revue (revues *gold* ou hybride) ou la politique des éditeurs en matière de dépôt en archive ouverte. L'IST peut aussi vérifier la politique de l'éditeur en matière de données, et d'éventuels liens entre un éditeur et un entrepôt.

Sélection et archivage

A l'échelle du projet, l'archiviste est le référent IST le mieux placé pour accompagner le chercheur afin de déterminer quelles sont les données à sélectionner pour être conservées, avec quelle durée d'archivage (pendant 10 ans, pour toujours...) et avec quelles solutions techniques en appui.

L'ensemble du DMP sera mis à jour en fonction de l'évolution du projet, par le chercheur qui en est responsable.

4. PERSPECTIVES

Et maintenant ?

Le modèle de DMP que nous avons présenté est public depuis quelques jours. Nous prévoyons de développer ce projet pour aller plus loin :

- élargir les coopérations autour de ce modèle de DMP avec les autres établissements de la ComUE Sorbonne Paris Cité, au travers notamment de sa Cellule Europe. Mais aussi travailler en lien avec des acteurs identifiés au sein de chacun des membres de la ComUE pour faciliter une appropriation du DMP par les établissements et une adaptation à leurs besoins spécifiques. Les premiers projets européens seront bientôt sélectionnés, la prochaine étape sera donc d'identifier clairement les acteurs ressources à SPC.

- faire le lien avec les groupes de travail au sein des associations professionnelles qui travaillent aussi sur la question mais sont pour le moment moins avancés dans l'approche pratique : l'ADBU (Association des Directeurs de Bibliothèques Universitaires) et l'AAF (Association des Archivistes Français).
- réfléchir à l'informatisation du DMP en mode agile, de façon à pouvoir le faire évoluer au fil de l'émergence de besoins et à intégrer des référentiels et des outils de workflow pour faciliter sa rédaction. L'ambition serait d'en faire un véritable outil de suivi moins chronophage pour les chercheurs et intégrant les compétences des différents acteurs.
- accentuer l'effort de formation et de sensibilisation en ouvrant plus largement les formations données pour les chercheurs et les ingénieurs-projets au sein de SPC, en enrichissant ces formations grâce au financement européen que nous avons obtenu dans le cadre du projet européen FOSTER (*Facilitate Open Science Training For European Research*).

ANNEXES

- Présentation powerpoint : A. Cartier, M. Moysan, N. Reymonet. *Construire des outils pour la gestion des données de la recherche dans une communauté d'universités* (16/01/2015)
- Document de référence : A. Cartier, M. Moysan, N. Reymonet. *Réaliser un plan de gestion de données : guide de rédaction* (V1, 09/01/2015)

REFERENCES BIBLIOGRAPHIQUES ET SITES DE REFERENCE

- Commission européenne. *Recommandation de la commission du 17.7.2012 relative à l'accès aux informations scientifiques et à leur conservation*. C(2012) 4890 final
http://medoanet.sciencesconf.org/conference/medoanet/pages/recommandation_access_and_preservation_scientific_information_fr_copie.pdf
- European Commission. *Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020* (dec. 2013)
http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf
- European Commission. *Guidelines on Data Management in Horizon 2020* (dec. 2013)
http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf
- European Commission. *Multi-beneficiary General Model Grant Agreement* (dec. 2013)
http://ec.europa.eu/research/participants/data/ref/h2020/mga/gga/h2020-mga-gga-multi_en.pdf
- ERC. *Open Access in Horizon 2020 - EC funded projects. Briefing paper for Project Coordinators*.
<https://www.openaire.eu/guide-for-project-coordinators-in-horizon-2020/document-details>
- OpenAire. <https://www.openaire.eu/guide-for-project-officers-in-horizon-202/view-document>
- OCDE. *Principes et lignes directrices de l'OCDE pour l'accès aux données de la recherche financée sur fonds publics* (2007) <http://www.oecd.org/fr/science/sci-tech/38500823.pdf>