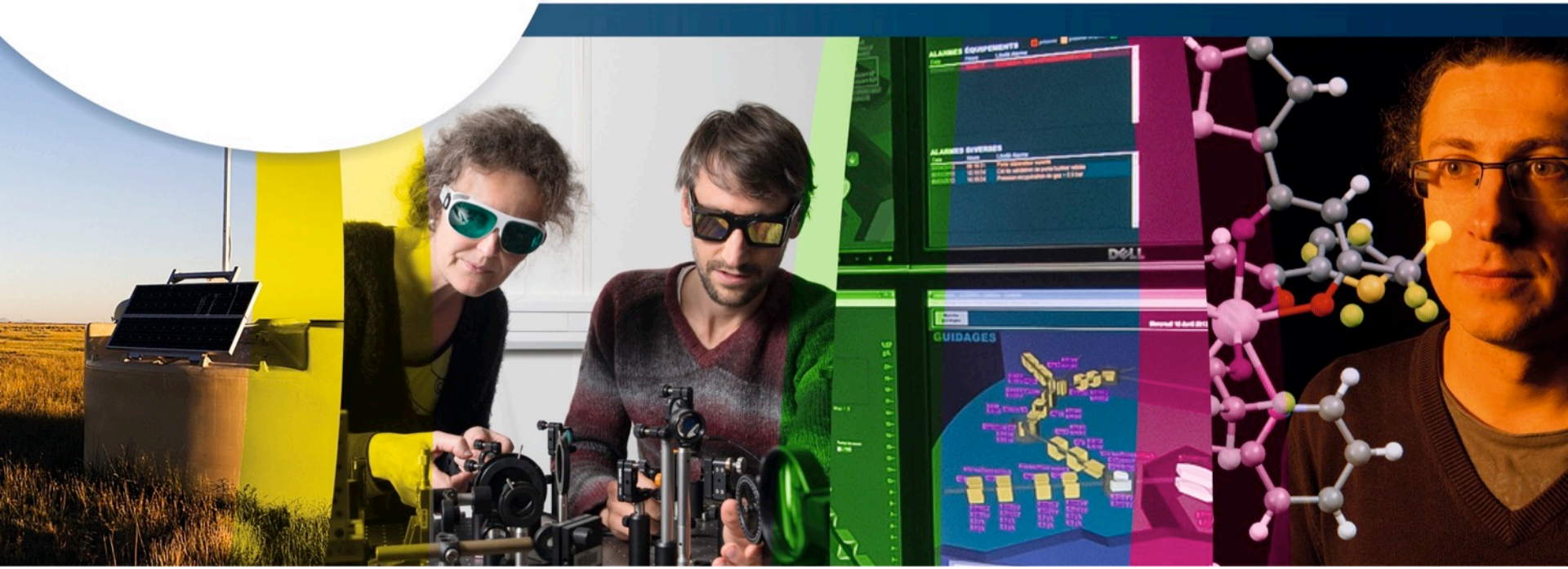
www.cnrs.fr

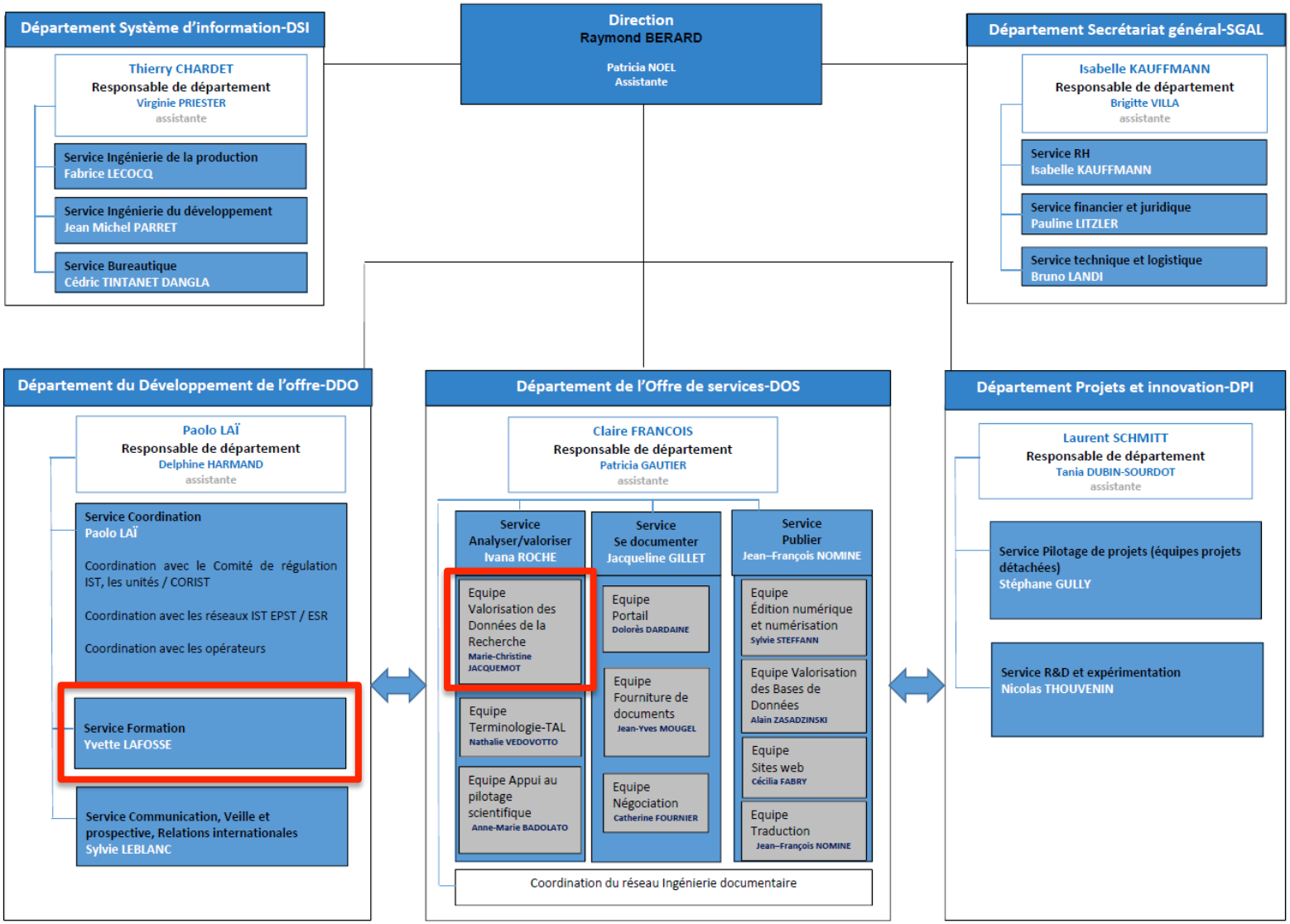
Accompagner les chercheurs : 5 tutoriels à votre disposition pour la rédaction d'un DMP dans le cadre d'Horizon 2020

-

Atelier FréDoc 2015



Organigramme de l'Inist-CNRS



Date de dernière mise à jour : 24/07/2015

www.cnrs.fr

Sommaire

- Contexte
- Objectifs
- Equipe projet
- Dispositifs de formation
- Entrée dans la pratique

  à partir des tutoriels sur le DMP

  - Questions à se poser
  - Bonnes pratiques de gestion
  - Outils

Contexte

29.3. Libre accès aux données de la recherche
[OPTION pour les actions participant au pilote ouvert sur les données de la recherche :
En ce qui concerne les données numériques de la recherche produites dans le cours de l'action (« **données** »), les bénéficiaires doivent :
Numéro de la convention de subvention: **[insérer le numéro] [insérer l'acronyme] [insérer le numéro d'identification de l'appel correspondant/principal/initial]**

(a) les déposer dans une banque de données de la recherche et prendre des mesures afin de permettre aux tiers d'accéder aux éléments suivants et de les explorer, exploiter, reproduire et diffuser, gratuitement pour l'utilisateur:

(i) les données, y compris les métadonnées, nécessaires pour valider dès que possible les résultats présentés dans des publications scientifiques;

*(ii) d'autres données, y compris les métadonnées associées, spécifiées dans le « **plan de gestion de données** » et dans les délais qui y sont fixés (voir annexe 1) ;*

(b) fournir des informations, par la banque de données, sur les outils et les instruments à la disposition des bénéficiaires et nécessaires pour la validation des résultats (et, si possible, fournir les outils et instruments eux-mêmes).

Cela est sans effet sur l'obligation de protéger les résultats énoncée à l'article 27, sur les obligations de confidentialité énoncées à l'article 36, sur les obligations en matière de sûreté énoncées à l'article 37 ou sur les obligations de protéger les données à caractère personnel énoncées à l'article 39, qui continuent de s'appliquer.

À titre exceptionnel, les bénéficiaires ne sont pas tenus d'assurer l'accès ouvert à des parties spécifiques de leurs données de recherche si la réalisation de l'objectif principal de l'action, tel que décrit à l'annexe 1, s'en trouvait menacée. En pareil cas, le plan de gestion des données doit indiquer les motifs de la non-accessibilité.

(Source : European Commission. Modèle général de convention de subvention multibénéficiaire. Version 1.0 [Internet]. 11 décembre 2013. [consulté le 18 août 2015]. Disponible sur : http://cache.media.education.gouv.fr/file/Affaires_juridiques_et_financieres/20/5/Convention_de_subvention_version_francaise_v1_376205.pdf

H2020 - article 29 du modèle de convention sur **le libre accès aux résultats de la recherche** (publications et données de recherche)

Objectifs

- Demande de la **Direction de l'Information Scientifique et Technique** (DIST) du **CNRS**

- **Accompagner et former les chercheurs**, enseignants-chercheurs, doctorants et les médiateurs : **Ingénieurs Projets Européens** (IPE), **correspondants Europe des Instituts**

- **Mettre en pratique les recommandations de la CE**

Equipe projet

- **DIST** : F André
- **Des partenaires externes associés** :
  - CCSD, Huma-Num, Universités Paris Diderot et Paris Descartes, INRA
- **Inist-CNRS : une équipe transversale de 11 personnes**, **spécialistes de la formation, du libre accès** (publications et données), **des données et de l'archivage**
- Un **soutien de la cellule données de recherche**
- Un **soutien du service Web**

Dispositifs de formation

- **Dispositif phase I – Formations à distance en différé (tutoriels)**
  - Se former où on veut, quand on veut et sur tout support
  - Durée courte : 10 mn par tutoriel maximum

- **D'autres dispositifs sont prévus en 2015-2016 :**
  - Dispositif phase II - Formations en présentiel à distance (classes virtuelles)
  - Dispositif phase III - Formations en présentiel à la demande
  - Intégration des tutoriels dans un LMS : gestion des apprenants et tutorat
  - Guichet numérique d'assistance et de conseil dédié (FAQ, Forum)

Dispositif phase I : 7 tutoriels

www.cnrs.fr

Un tutoriel sur les principes généraux de dépôt des publications et données de la recherche dans le cadre d'Horizon 2020

Cinq tutoriels sur la rédaction d'un DMP (Data Management Plan), c'est-à-dire plan de gestion des données

Un tutoriel sur le dépôt des données de recherche dans un entrepôt

Perspectives / Planning

- ***Dispositif phase I* – Formations à distance en différé (tutoriels)**
  - **Améliorations et mises à jour**
  - **Conception de nouveaux tutoriels :**
    - **Identifiants pérennes, DOI, DataCite, dépôt dans HAL** (en collaboration avec le CCSD)…
    - **En fonction des demandes**
- ***Dispositif phase II* - Formations en présentiel à distance (classes virtuelles)**
- ***Dispositif phase III* - Formations en présentiel (face à face) à la demande**
- ***Intégration des tutoriels dans un LMS : suivi des apprenants et tutorat, signalement dans FUN (*** ***http://www.france-universite-numerique.fr/moteur-ressources)***
- ***Guichet numérique d'assistance-conseil* (FAQ, Forum)**

→ ***fin 2015 – 2016***

Projets de partenariat / Mutualisation

- Inist - CNRS / Urfist / Renatis financé par le MENESR :
  - Projet de formations à distance sur la gestion, la préservation, le partage et la diffusion en libre accès des données de la recherche

- Renatis / Inist-CNRS / Urfist :
  - ANF « Participer à l'organisation du management des données de la recherche : gestion de contenu et documentation des données » (en cours de validation)

Sondage

- Connaissez-vous ces tutoriels ?
- Avez-vous déjà accompagné des chercheurs
  - dans la rédaction d'un DMP ?
  - dans la gestion et la diffusion de leurs données de recherche ?
- Quels outils ou ressources avez-vous utilisés ?

Tutoriel DMP – Présentation générale

- Structuration d'un DMP
- Utilité d'un DMP
- Rédaction d'un DMP
- Outils

http://www.inist.fr/formations/DMP-Presentation-generale/story.html

Activité pratique

- Exemple de projet de recherche
  - Titre : Rôle de la température sur le dépérissement des aulnes dans le Nord-est de la France (PHYTOCLIM)*
  - Description et méthodologie
    - Sites géolocalisés sur un gradient de température (données Météo-France)
    - Repérage et suivi de l'état sanitaire de 50 aulnes/site
    - 2 barèmes pour chaque arbre
      - Présence ou absence de chancre à la base du tronc
      - Niveau de dépérissement de la cime des arbres (0 à 4)

*Aguayo J, Elegbede F, Husson C, Saintonge FX, Marçais B. Modeling climate impact on an emerging disease, the Phytophthora alni-induced alder decline.. Glob Chang Biol. 2014;20:3209-21

Activité pratique

Tableau 1 : Caractéristiques des sites

Numéro du site	Longitude en WGS84	Latitude en WGS84	Nom de la rivière	Station météo
1	49,1	6,2	Moselle	Metz
2				
…				
16				

Activité pratique

Tableau 2 : Notation

Numéro du site	Arbre	Année notation	Chancre	Cime	Temp moy hivernale (degré)	Temp moy estivale (degré)
1	1	2007	0	1	4,1	19,1
1	2	2007	1	2	4,1	19,1
1	3	2007	0	2	4,1	19,1
1	4	2007	0	4	4,1	19,1
1	…	2007	0	3	4,1	19,1
1	50	2007	1	2	4,1	19,1
1	1	2008	1	1	3,5	17,8
1	2	2008	1	1	3,5	17,8
1	3	2008	0	2	3,5	17,8

DMP – Modèle de la Commission européenne

- **Référence et nom du jeu de données**

Identifiant du jeu de données qui va être produit.

- **Description du jeu de données**

Celle-ci décrit les données qui seront générées ou collectées, leur origine (s'il y a collecte), leur nature et leur échelle, à qui elles pourraient être utiles et si elles viennent en appui d'une publication scientifique, mais aussi les informations sur l'existence (ou la non-existence) de données similaires et leurs possibles intégrations et réutilisations.

- **Normes et métadonnées**

Indication des normes existantes relatives à la discipline. À défaut, un résumé expliquant quelles métadonnées seront créées et comment.

- **Partage des données**

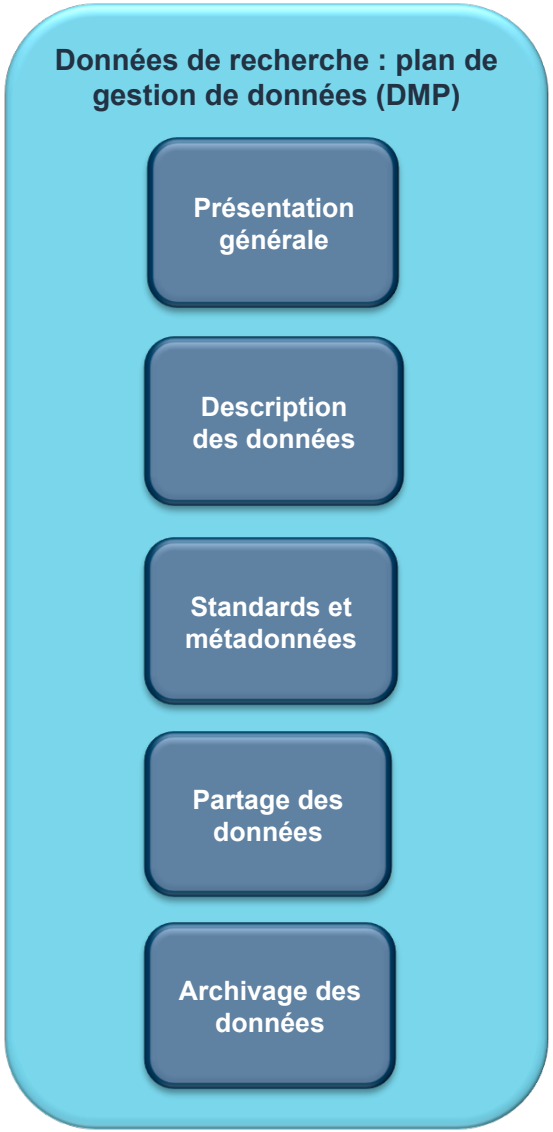
Description des modalités de partage, y compris les procédures d'accès, les périodes d'embargo (le cas échéant), exposé succinct des dispositifs techniques de diffusion et logiciels nécessaires ainsi que les autres outils permettant la réutilisation de données, et indications sur le caractère librement accessible des données ou si celles-ci seront limitées à des groupes spécifiques. Identification de l'entrepôt dans lequel les données seront stockées, s'il existe déjà et est identifié, en indiquant son type (institutionnel, ou de référence dans la discipline considérée, etc.).

Les raisons d'une éventuelle impossibilité de partage d'un jeu de données devront être indiquées (par ex. motifs de nature éthique, liés à la protection des données personnelles, de la propriété intellectuelle, commerciale, à la protection de la vie privée, et de sécurité, etc.).

- **Archivage et conservation (y compris stockage et sauvegarde des données)**

Description des procédures mises en place pour une conservation à long terme des données. Indication de la durée de conservation souhaitable des données, leur volume approximatif, les coûts associés et la façon donc ces coûts seront couverts.

Source : Inist-CNRS Lignes directrices pour le libre accès aux publications scientifiques et aux données de recherche dans Horizon 2020. Traduction française non officielle [Internet]. 19 septembre 2014. [consulté le 19 août 2015]. Disponible sur : http://openaccess.inist.fr/?Traduction-francaise-des-deux



A vous de jouer !

- A partir du modèle et des différents tutoriels
  - Accompagnement d'un chercheur dans la rédaction d'un DMP
  - Des bonnes pratiques et outils/ressources à connaître

Tutoriel DMP - Description des données

- Origine des données
  - Création ou données existantes – Méthodologie
  - Type de données (source, forme, stabilité, volumétrie)
- Description formelle et technique
  - Format, ouverture, encodage, compression
  - Système d'organisation, de nommage et de gestion des répertoires et des fichiers

DMP - Description des données

Description du jeu de données
(modèle de la Commission européenne)

« Celle-ci décrit les données qui seront générées ou collectées, leur origine (s'il y a collecte), leur nature et leur échelle, à qui elles pourraient être utiles et si elles viennent en appui d'une publication scientifique, mais aussi les informations sur l'existence (ou la non-existence) de données similaires et leurs possibles intégrations et réutilisations »

- **Types de données ?**
- **Formats de données ?**
- **Volumétrie** ?

Tutoriel DMP - Standards et métadonnées

- A quoi servent les métadonnées ?
- DMP et métadonnées
- Comment choisir les standards de métadonnées ?
- Un exemple de métadonnées
- Comment créer les métadonnées ?

DMP - Standards et métadonnées

Normes et métadonnées

(modèle de la Commission européenne)

« Indication des normes existantes relatives à la discipline. À défaut, un résumé expliquant quelles métadonnées seront créées et comment. »

www.cnrs.fr

Comment choisir les standards de métadonnées ?

- Critères de choix ?
- Ressources
  - DCC http://www.dcc.ac.uk/resources/metadata-standards
  - RDA Metadata Standards Directory Working Group. Metadata Repository. http://rd-alliance.github.io/metadata-directory

Tutoriel DMP - Partage des données

- Pour chaque jeu de données,
  - Les modalités de dépôt
  - Les modalités de partage
  - Les licences
  - Les exceptions

DMP - Partage des données

Partage des données (modèle de la Commission européenne)

« Description des modalités de partage, y compris les procédures d'accès, les périodes d'embargo (le cas échéant), exposé succinct des dispositifs techniques de diffusion et logiciels nécessaires ainsi que les autres outils permettant la réutilisation de données, et indications sur le caractère librement accessible des données ou si celles-ci seront limitées à des groupes spécifiques. Identification de l'entrepôt dans lequel les données seront stockées, s'il existe déjà et est identifié, en indiquant son type (institutionnel, ou de référence dans la discipline considérée, etc.).

Les raisons d'une éventuelle impossibilité de partage d'un jeu de données devront être indiquées (par ex. motifs de nature éthique, liés à la protection des données personnelles, de la propriété intellectuelle, commerciale, à la protection de la vie privée, et de sécurité, etc.). »

- **Identifier un entrepôt de données de la recherche ?**
  - http://www.re3data.org/
- **Exceptions au partage des données ?**
- **Licences Creative Commons ou autres ?**
- **Fournir les outils** nécessaires à la validation/réutilisation des données
- **Identifiants pérennes**

Tutoriel DMP - Archivage des données

- Qu’est-ce que l’archivage des données
- Comment réaliser un archivage des données ?
- Où s’adresser ?

DMP - Archivage des données

Archivage et conservation (y compris stockage et sauvegarde des données) (modèle de la Commission européenne)

« Description des procédures mises en place pour une conservation à long terme des données. Indication de la durée de conservation souhaitable des données, leur volume approximatif, les coûts associés et la façon donc ces coûts seront couverts. »

- **Sauvegarde des données ?**
- **Conservation à long terme des données?**

Webographie
Plan de gestion des données

- Cartier A, Moysan M, Reymonet N (2015). Réaliser un plan de gestion de données. Université Paris Diderot et Université Paris Descartes, 30 p. http://www.univ-paris-diderot.fr/DocumentsFCK/recherche/Realiser_un_DMP_V1.pdf
- Deboin MC (2014). Découvrir des plans de gestion des données de la recherche, en 4 points. Montpellier (FRA) : CIRAD, 6 p. http://coop-ist.cirad.fr/content/download/5435/40362/version/4/file/CoopIST-plan-gestion-donnees-recherche-20140717.pdf
- Digital Curation Centre (2013). Checklist for a Data Management Plan. v.4.0. Edinburgh: Digital Curation Centre. http://www.dcc.ac.uk/resources/data-management-plans
- DMPonline. Edinburgh : Digital Curation Centre. 2004-2015 [consulté le 10 septembre 2015]. https://dmponline.dcc.ac.uk

Webographie
Plan de gestion des données (suite)

- DMPTool. Data Mananagement Planning Tool. University of California Curation Center of of the California Digital Library. 2010-2015 [consulté le 10 septembre 2015]. https://dmptool.org
- FOSTER - Facilitate Open Science Training For European Research. Research Data Management Plans. https://www.fosteropenscience.eu/taxonomy/term/139
- Romier G (2014). Une préoccupation partagée : plan de gestion des données et projets Horizon 2020 (France Grilles). In : RBDD - Journée de sensibilisation à la sécurisation et à la pérennisation des données [en ligne]. Institut d'Astrophysique de Paris. [Consulté le 09 septembre 2015]. Disponible à l'adresse : http://rbdd.cnrs.fr/IMG/pdf/francegrilles-rbdd6nov2014v2.pdf?104/d9e53228e07433f0a0f081a31cf232b100b0c050

www.cnrs.fr

Webographie - Pour l'atelier

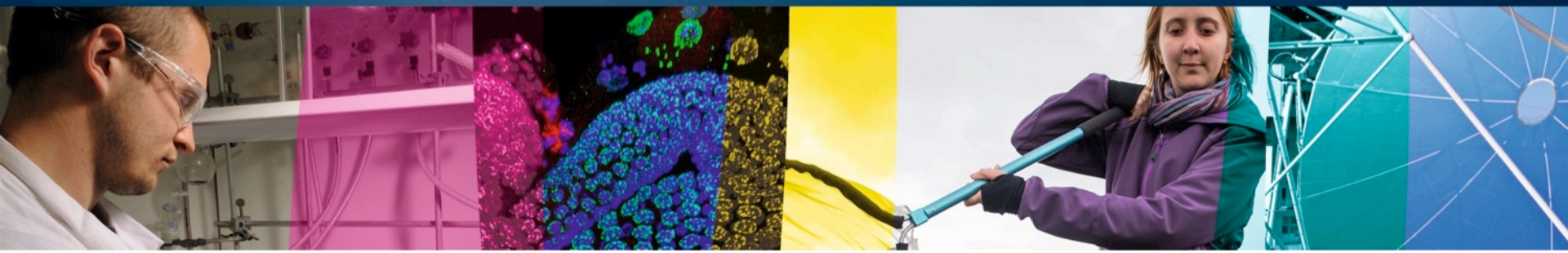
- Description du jeu de données (formats)
  - DMPTool - Data Management General Guidance : https://dmptool.org/dm_guidance
  - UK Data Archive – File Formats Table : http://www.data-archive.ac.uk/create-manage/format/formats-table
  - Purdue University Research Repository - File Format Recommendations : https://purr.purdue.edu/legal/file-format-recommendations

Webographie - Pour l'atelier (suite)

- Normes et métadonnées
  - DCC - Disciplinary Metadata : http://www.dcc.ac.uk/resources/metadata-standards
  - RDA Metadata Standards Directory Working Group. Metadata Repository : http://rd-alliance.github.io/metadata-directory
- Partage des données
  - re3data : http://www.re3data.org
  - Licences Creative Commons : http://creativecommons.org/choose

www.cnrs.fr

Contact

Pour des questions ou des compléments d'informations :

Service Formation
contact-formation@inist.fr
03 83 50 94 04
Equipe Valorisation des Données de la Recherche
Marie-Christine.jacquemot@inist.fr