



LA GESTION DES DONNEES DE LA RECHERCHE

Séminaire du LEST

Aix-en-Provence le 12 septembre 2016

Anne Bonneville, Service documentation du LEST



PLAN

1. Données de la recherche : essai de définition
2. Contexte institutionnel et de la Recherche
3. Enjeux
4. Quelques recommandations
5. Quelques liens ...

1. DONNÉES DE LA RECHERCHE : ESSAI DE DÉFINITION

- Définition de l'OCDE (2007) pour les données financées sur Fonds publics : « Les données de la recherche sont définies comme des *enregistrements factuels (chiffres, textes, images, sons)* utilisés comme source principale pour la recherche scientifique et généralement reconnus par la communauté scientifique comme nécessaires *pour valider les résultats de recherche.* »
- Définition institutionnelle
« Toutes les données de recherche (natives ou non) produites dans un processus de recherche, financées sur Fonds publics ou réutilisées pour la recherche
Plusieurs composantes :
 - les données primaires/brutes issues de l'observation, uniques/contexte
 - les données traitées ou dérivées (ex données de simulation - modèles éco)
 - les données analysées, interprétées (ex : Tables/Base de données)

CONTEXTE POLITIQUE ET INSTITUTIONNEL

- 2002 : Initiative de Budapest (Open access)
- 2003 : Déclaration de Berlin sur le « Libre accès à la connaissance scientif. (CNRS...)»
- 2004 : Déclaration de l'OCDE sur l'accès aux données de la recherche financée par des Fonds publics (France)
- 2009 : Projet national - BSN 10 (MESR) sur la gestion des données

- 2013 : EUROPE
 - > Recommandation de la ligue européenne des universités de recherche (LERU) qui préconise une politique institutionnelle sur les « research data »
 - > Open Research data Pilot - Projet du programme H2020 (CEE) + Zenodo

- : FRANCE
 - > TGIR Huma-Num : Données de recherche de toute nature en SHS
 - Progedo : Données quantitatives en sciences sociales

- 2014-... : Plan d'actions de la DIST (CNRS)
Groupe de travail (3.2) dédié aux données de la recherche »

Des initiatives sur le terrain d'offres de services, et d'expertises dans certaines universités et EPST (INIST CNRS, IRSTEA, SCD Univ. Lille ; SCD Univ. Paris Diderot, SCD Univ. Nice, ...)

2. CONTEXTE DE LA RECHERCHE

- Nouveau paradigme (« open science ») / Nouvelle façon de faire la science : plus ouverte, collaborative, numérique et citoyenne
- Explosion et fragilité des données numériques
- Invisibilité des données produites
- Evolution des modes de publication:
 - Publication de données**
 - Fichiers joints à la publication pour les données (supplement materials)
 - Liens vers des données déposées sur un entrepôt (ex : IJESRM/Zenodo)
 - Nouvelles revues dédiées à la description des données : « Data Journal »
Ex : Journal of Open Archeology Data
 - Politiques chez certains éditeurs :**
Ex : revues concernées en Sciences sociales
Ex : American Economic Review (*un standard en économie*)

- De nouveaux outils de recherche : Ex : Data search (Elsevier) ; Re3Data

ENJEUX

- **Enjeux éthiques**

- Qualité des données
- Confidentialité des données (*données sensibles*)
- Transparence, Intégrité de la recherche (*reproduction, validation des résultats*)

- **Enjeux économiques**

- Financement des recherches publics (*retour sur investissement/doublons*)
- Marché de la donnée
- Citation des données

- **Enjeux pour le chercheur**

- Conservation à long terme de ses données (*à valeur potentielle*)
- Réutilisation des données par le déposant et ses pairs (*iD/URL pérenne, licence*)
- Transparence, fiabilité des données réutilisées
- Gain de temps dans une nouvelle recherche
- Lien bidirectionnel entre les jeux de données et les articles publiés
- Citation des données
- Visibilité accrue

QUELQUES RECOMMANDATIONS

- « Pensez : « Données » dès le démarrage d'une recherche et tout au long du cycle du projet
- Plan de gestion des données dans le cadre de vos contrats financés :
contexte, responsabilité, types de données collectées ou produites, leurs descriptions (meta-données), arborescence de classement ; sécurité des données ; archivage à long terme et partage des données (*entrepôt de confiance, data journal, licence de réutilisation...*)
- **Bonnes pratiques** : règles de nommage des fichiers et formats recommandés, classification des fichiers, codification des versions, sécurité (cryptage des données sensibles) et sauvegarde des données sur un entrepôt public, attribution de licence d'utilisation pour protéger les données...

QUELQUES LIENS ...

- [Sauvegarder les données de la recherche](#) (Blog Présidence de Lille)
- [Gestion des données de la recherche : quelques pistes pour démarrer](#) (INSHS)
- [Gestion et diffusion des données de la recherche](#) (URFIST Rennes)
- [Réaliser un Plan de gestion des données](#) (Paris Descartes)
- [Journée de formation sur le management des données de la recherche](#) (ANF CNRS 2016)
- [DoRANum](#) (BSN10) : Ressources d'autoformation sur la gestion et le partage des données
- [Guide « form@doct »](#) (Université de Bretagne)