

LES ENJEUX DE L'OPEN ACCESS (2) L'ARCHIVAGE / LA MISE EN LIGNE DES DONNEES DE LA RECHERCHE



La notion de données de recherche

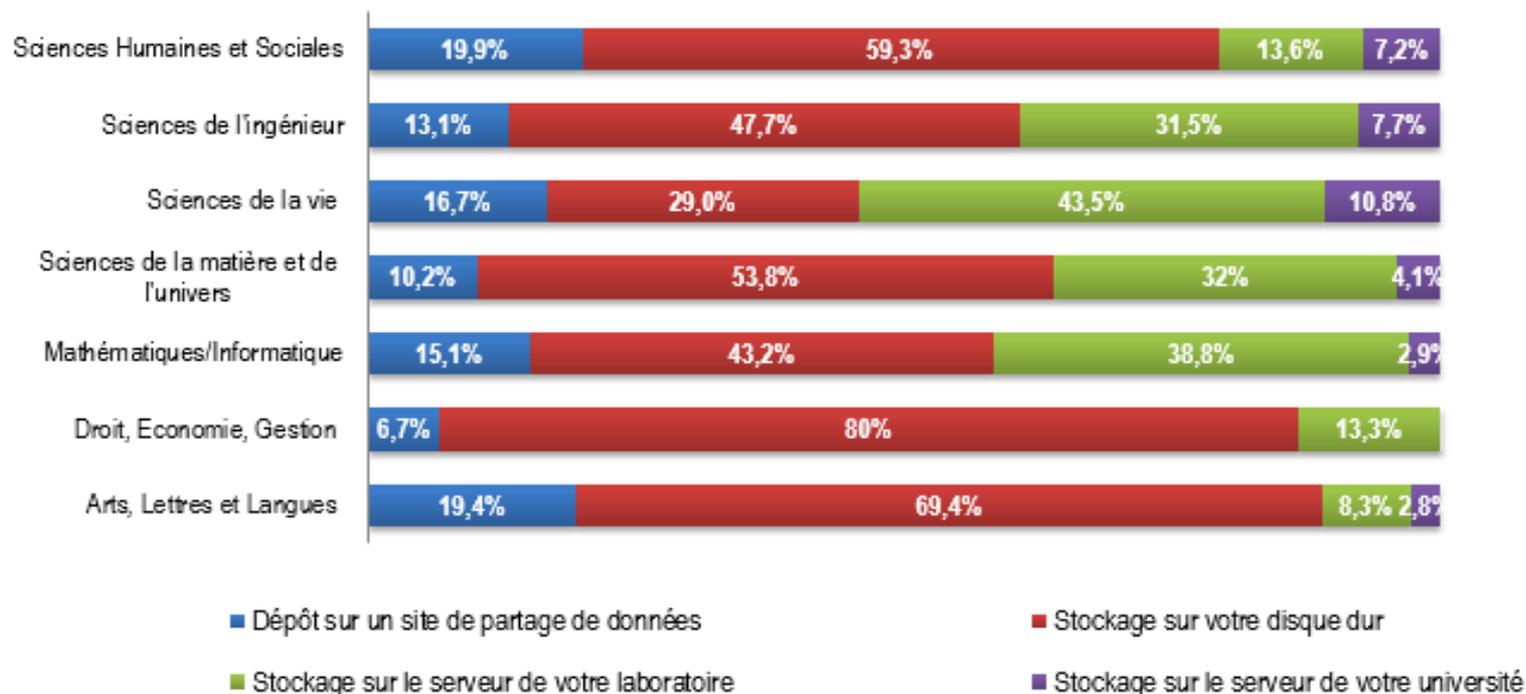
- Définition :
 - Tous types de données produites lors d'une recherche résultant d'observations, d'enquêtes, d'expériences (datasets).
 - Ces données sont une source à partir de laquelle sont produites les publications

Gérer les données de recherche ? Pendant un projet

- Organiser la conservation des données de recherche tout au long d'un projet :
 - Créer des dossiers et des noms de fichiers significatifs
 - Conserver les versions pour documenter la progression de l'étude
 - Organiser cette conservation dans un lieu informatiquement sûr : sur un serveur d'université par ex. et non sur un disque dur

Extrait Enquête Couperin 2014

Stockage des données de recherche

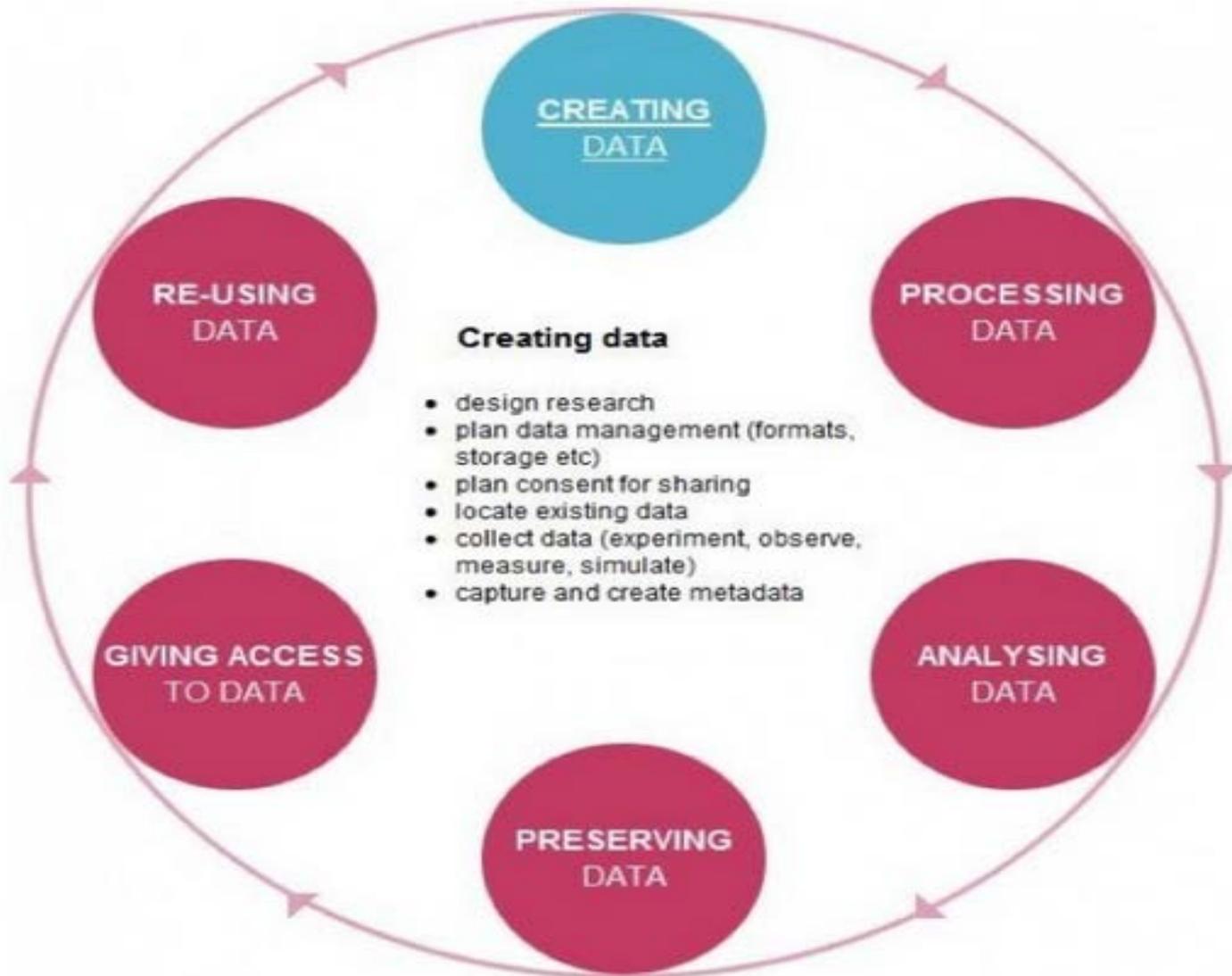


40% des chercheurs n'utilisent que leur disque dur et souhaiteraient que leur université leur fournisse une infrastructure logistique préservant leurs données de recherche

- Décrire les données
 - Renseigner les métadonnées décrivant les données pour pouvoir les interpréter à court et long terme
 - Signaler les logiciels spécifiques permettant de les lire, si nécessaire
 - Anonymiser les données si nécessaire - contraintes juridiques
 - Pouvoir plus facilement les trouver

- **Rendre accessibles les données à l'appui des résultats décrits dans une publication**
 - Ethique de la recherche : tous les résultats doivent être démontrables, reproductibles, « critiquables »
 - Accessibilité des données étayant un article : exigence de certaines revues de recherche
 - Production de recherche à part entière : les munir d'un identifiant (DOI) pour les rendre citables

Data Life Cycle



L'incitation européenne

- Programme H2020 européen stipule que les bénéficiaires du financement de leur projet de recherche (2016-2017) :
 - “must deposit in a research data repository and take measures to make it possible for third parties to access, mine, exploit, reproduce and disseminate - free of charge for any user - the following:

- (i) the data, including associated metadata, needed to validate the results presented in scientific publications as soon as possible;
- (ii) other data, including associated metadata, as specified and within the deadlines laid down in the data management plan.” (EC Guidelines, p.3.)

- Est une obligation pour certains domaines couverts par le programme :
 - technologies émergentes, villes intelligentes, infrastructure de la recherche, TIC, nanotechnologies, défis sociétaux (climat, développement durable), sociétés innovantes
- Est possible sur la base du volontariat pour les autres domaines

- Possibilité de « opt-out » si : pas de données, protection des résultats, confidentialité, sécurité, données personnelles
- Possibilité d'ouvrir une partie des données
- Sur 3700 contrats signés, 35% ont choisi opt-out (raisons juridiques ...), 12% opt-in volontaire

- Les Plans de gestion des données (DMP)
 - Prise en compte du cycle entier de vie de la donnée : quels standards utilisés, que préserver, que partager ?
 - Adaptation au domaine de recherche à effectuer
 - Collaboration entre chercheur (détenteur des données) / Service informatique (infrastructure) / bibliothèque (description, diffusion) / archives (conservation)

- Exemple de Data Management Plan :
Paris 5 / Paris 7 : http://www.univ-paris-diderot.fr/sc/site.php?bc=recherche&np=Data_DM&g=m

- Ressources générales :

- Digital Curation Centre : <http://www.dcc.ac.uk/>

- Rubrique sur les données de recherche au European University Institute

- <http://www.eui.eu/Projects/EUIDataRepository/Home.aspx>

Les données de recherche en Open Access ?

- Pourquoi ?
 - Ethique de la recherche : à l'appui des articles mis en OA, pour une recherche vérifiable, transparente et efficace
 - Rendre les données utilisables pour développer l'innovation
 - Même philosophie que pour les articles : éviter que seuls les journaux commerciaux disposent de ces données et fassent payer les accès

- Comment partager des données ?
 - Entrepôts thématiques existent pour partager des données mais en économie aucun ne s'impose (il n'y a pas de RePEc des données)
 - Archives institutionnelles. Ex. Gutteridge, Christopher (2010) Using the Institutional Repository to publish research data. In, *Open Knowledge Conference (OKCon) 2010, London, UK*,
<http://eprints.soton.ac.uk/id/eprint/270885>

Trouver des données de recherche

- Catalogue international d'entrepôts de données de recherche : Registry of Research Data Repositories re3data.org
 - 116 entrepôts donnant accès à des données en Economie
 - Ex en sciences sociales : [Centre de données socio-politiques \(CSDP\)](#) de Sciences Po a pour mission de participer au développement et à l'amélioration des méthodes qualitatives des sciences sociales en France, notamment par la diffusion de documents contextualisés et analysés. Le CSDP fait partie du réseau européen de données en sciences sociales CESSDA (Council of European Social Science Data Archives).

