

# Données d'enquêtes

## Exemples à la bibliothèque

La gestion des données de la recherche dans la valorisation  
de la production scientifique : l'offre de service des bibliothèques

*MESRI-ADBU / 21 novembre 2017*

Sophie Forcadell, Appui à la recherche, DRIS, Sciences Po

## Éléments de contexte

### Enjeux pour les chercheurs

- Garder la maîtrise des données, les réutiliser, les partager

### Enjeux pour l'institution

- Assurer un environnement favorable au développement de la recherche
- Rendre visible et valoriser la production scientifique
  - Être acteur d'une dynamique nationale en SHS et s'inscrire dans un écosystème international

2015-2016 : Exploration et préconisations

2017-2018 : Expérimentation "grandeur nature" du programme

## Sciences Po, une histoire de données

### Centre de données socio-politique (CDSP - 2006)

- Expertise données enquêtes qualitatives et quantitatives depuis +10 ans
- Projets nationaux structurants : [DIME-SHS](#) ([BeQuali](#), Panel [Elipss](#), [Hyphe](#))

### Centres de recherche et projets

- [Archipolis](#) (2012)
- [médialab](#) (2009)
- [METSEM](#) (2016)

## Activités expérimentées par la DRIS

### Organiser

- Conseiller les chercheurs pour la planification de la gestion des données (Data Management Plan) dès le montage d'un projet

### Expertiser

- Participer à la préparation, à la sélection, au recueil, au traitement, au dépôt, à la documentation et à l'archivage des données

### Sensibiliser et former

- Développer les bonnes pratiques et l'usage des normes et règles méthodologiques, techniques et légales.

## Outil : Entrepôt DataVerse

- Orienté sur les objets de sciences sociales
- Solution Harvard développée par des chercheurs de l'Institute for Quantitative Social Science (IQSS)
- **Capitalisation de l'expérience** du CDSP et du consortium Archipolis
- Coût d'entrée faible **en infrastructure** pour la DRIS
- Permet l'expérimentation sur les projets pilotes
- Permettra ensuite d'exposer les données (visibilité, catalogue)
- Schéma de métadonnées DDI adapté aux types de données produites à Sciences Po
- Ajout automatique d'un DOI pour chaque jeu de données
- Interopérabilité avec l'archive institutionnelle SPIRE (passerelles à développer)
- Complémentarité avec GitHub **si développement de logiciels ad-hoc en parallèle**

## Projets pilotes : Cas 1

- **Projet européen H2020** en cours mené dans plusieurs grandes villes européennes (Copenhague, Berlin, Paris, Londres etc.) sur les politiques publiques en matière de transport. Equipe internationale.
- **Enjeux juridique et technique** :
  - la chercheuse souhaite partager et aussi réutiliser les données mais doit résoudre les questions juridiques et éthiques.
  - données nombreuses (+200 fichiers) et hétérogènes (rapports, données tabulées de formats divers, notes, fichiers audio, transcriptions), données primaires et secondaires.
- Travail en cours : Récolement ; ré-encodage ; conversion dans des formats ouverts et pérennes
  - Point de vigilance sur la sélection des fichiers et l'organisation des dossiers pour la collecte, la réutilisation et le partage
- Accompagnement dans de **nouvelles méthodes de travail** vs “vous devez faire cela”

## Projets pilotes : Cas 2

- Projet achevé, objet d'une **valorisation des données à travers un corpus outillé** (site web avec possibilité d'explorer/de jouer les données). Données d'échanges commerciaux bilatéraux.
- Les développements, interface et données en .csv sont diffusées sur GitHub, ce qui nécessite une reprise de la documentation et dépôt sur Dataverse.
- Points forts :
  - Données “propres” car BDD constituée avec le concours du médialab et en partie documentée (variables, processus de constitution de la base et traitement des données). Existence d'un datapaper.
  - Equipe consciente des enjeux de la diffusion et de la complémentarité de la bibliothèque dans le processus.
- Travaux en cours : Récupération des fichiers (+30) et récolement ; Vérification des formats, encodages et nommages des fichiers ; Documentation des jeux de données avec des métadonnées descriptives, de structure et administratives ; Dépôt des jeux de données sur le DataVerse.

## Projets pilotes : Cas 3

- Pas de projet !
- A l'origine : **une demande d'accompagnement** d'une chercheuse sur la gestion de photographies d'archives. Comment documenter les métadonnées images (IPTC) ?
- La demande formulée a trouvé écho chez d'autres chercheurs
- **Travail collectif** : bibliothèque (+ iconographe de la bibliothèque) ; ingénieur d'étude dans l'unité de recherche ; chercheurs
- Travail en cours : **organisation** de cette communauté ; **formation** aux métadonnées IPTC ; recommandation d'un logiciel de traitement ; préconisations de traitement



## **Données de la recherche / DRIS**

Contact : Cynthia Pedroja, chef de projet, DRIS, Sciences Po

<http://www.sciencespo.fr/bibliotheque/fr/enseignants-chercheurs/gestion-donnees-recherche>