

# Définition(s) des données de la recherche

Dans un DMP, il faut décrire les données de recherche du projet et décrire la manière dont elles seront gérées. Avant d'entamer ce travail, il est bon de se demander : qu'est-ce qu'une donnée de recherche ? Et quelles sont les données de recherche de mon projet ?

## Définition(s) des données de recherche

### Données de la recherche : quésaco ?

- **Des éléments factuels** : chiffres, textes, images, sons, mesures, résultats d'enregistrements, de programmes informatiques...
- **Bruts** (c'est-à-dire qui n'ont fait l'objet d'aucun traitement, d'aucune manipulation, d'aucune transformation) ou **dérivés de données brutes** (c'est-à-dire obtenus après transformation de données brutes)
- **De nature quantitative** (chiffres, mesures, statistiques, réponses à des enquêtes) **ou qualitative** (réponses à un entretien, discours, enregistrements de locuteurs, vidéos)
- **Sur lesquels sont basés la recherche**
- **Nécessaires selon la communauté scientifique pour valider les résultats de la recherche**
- **Qui peuvent être conservés sur tout support** (papier, numérique,...) **et format** (.png, .mpeg, .svg, .wma, .pdf, .txt, .xml...)

### Quelques points-clés pour mieux comprendre et identifier les données de la recherche

1. **A la base de la recherche** : Pour faire son travail, un chercheur a besoin de **données de recherche**. Elles sont **à la base** de son travail. Sans données, aucun résultat de recherche ne peut être testé, ni vérifié.



2. **Factuelles** : Les données de la recherche sont des **éléments factuels**. Par exemple, il peut s'agir de chiffres, de textes, d'images, de sons, de mesures, de résultats d'enregistrements, de résultats de programmes informatiques...
3. **Brutes ou dérivées** : Les données de la recherche peuvent être **brutes**, c'est-à-dire qu'elles n'ont pas encore été traitées, interprétées ou modifiées.

Elles peuvent aussi être **dérivées de sources brutes**, c'est-à-dire qu'elles correspondent à des données brutes transformées (traitées, organisées, nettoyées...).

4. **Nécessaires** : Les données de la recherche peuvent être de **nature** très **différente selon la discipline scientifique**. En général, un chercheur peut identifier ses données de recherche en se basant sur **celles** qui sont **reconnues dans sa discipline** comme nécessaires pour valider les résultats de la recherche.

Par exemple :

- en **archéologie**, les données de la recherche peuvent être des photographies, des relevés topographiques, des mesures dendrochronologiques, un SIG, une datation radiocarbone...
- en **sociologie**, il peut s'agir de séries statistiques, de résultats d'enquêtes, de sondages, d'images...
- en **sciences de la Terre et de l'environnement**, on peut utiliser des cartes géologiques, des photographies, des relevés de température, des calculs du taux de carbone présent dans la mer, des coordonnées GPS...

5. **Quantitatives ou qualitatives** : Les données de recherche peuvent être **quantitatives** (données chiffrées) ou **qualitatives** (données non-chiffrées).
6. **Conservables sur tout support et format** : **En théorie**, les données de la recherche peuvent être conservées sur tout support (papier ou numérique) et tout format (.pdf, .txt, .png, .xml...).



**Dans les faits**, lorsque l'on se place dans le contexte de l'open science et de l'open data, les **données** de recherche sont **souvent numérisées ou numérisables** afin de rendre leur stockage et leur partage possible et plus efficace.

**Au moment de choisir** le format d'enregistrement de ses données, il est vivement recommandé de respecter le principe : "Aussi ouvert que possible, aussi fermé que nécessaire" et de **choisir des formats dits "ouverts", non-propriétaires**. Par exemple, pour un fichier audio, mieux vaut privilégier le format .wav au format .mp3.

En effet, un fichier audio conservé :

- au format .MP3 ne pourra pas être lu par tous les logiciels capables de lire des fichiers audio : il s'agit d'un format fermé.
- au format .WAV pourra être créé, lu ou modifié par tous les logiciels destinés à lire des fichiers de type audio : il s'agit d'un format ouvert.

#### 7. **En bref :**

Les données de la recherche sont donc tous les éléments factuels qu'un chercheur peut **produire** ou **collecter** et dont il a besoin pour **valider les résultats** de sa recherche.

Il est important de noter qu'il existe plusieurs définitions des données de la recherche. Elles restent une notion **difficile à définir** car elles englobent un panel diversifié de réalités. D'une discipline à une autre, ou d'un point de vue institutionnel à un autre, la définition des "données de la recherche" est plus ou moins restrictive.

