



DATA  
INFRASTRUCTURE

*Développer  
la culture des données*



L'Adisp-Progedo : un portail de stockage et de diffusion pour les données historiques.

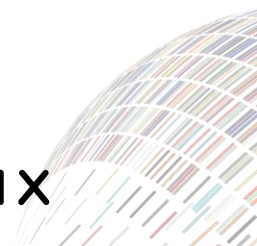
Le 25 novembre 2021

Alena Lapatniova (Assistante d'Ingénieur à l'ADISP - Progedo)



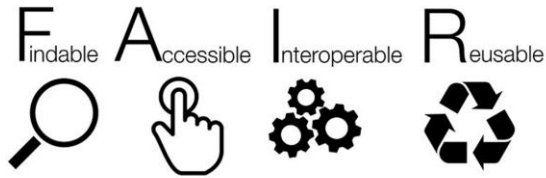
# De la statistique publique aux données de la recherche

- Depuis plus de 20 ans l'Adisp documente et diffuse aux chercheurs SHS des données issues de la statistique publique.
- Aujourd'hui, dans le contexte de la science ouverte les chercheurs ont l'obligation de partager leurs données à l'issue d'un projet de recherche financé par des fonds publics, tout en respectant les principes du RGPD.
- Désormais, l'Adisp propose son service de documentation et de diffusion (FAIR depuis l'origine et en respect du RGPD) à la communauté scientifique et, en particulier, aux



## Les AVANTAGES à diffuser les données historiques par PROGEDO

1er avantage : l'Adisp garantie la description fine des données (métadonnées) en utilisant le standard international DDI ce qui rend les données FAIR.



Démarche « FAIR »  
(Faciles à trouver, Accessibles,  
Interopérable et Réutilisables)

Concrètement, le schéma de description des données DDI permet :

- Décrire le projet de recherche dans le cadre duquel les données ont été collectées, préciser ses dimensions géographiques et temporelles et la population faisant l'objet de l'étude.
- Décrire les éléments méthodologiques tel que le plan et mode de collecte des données, la méthode d'échantillonnage, les sources primaires et secondaires utilisées, leur histoire et le lieu de leur conservation.
- Joindre les liens des publications apparentées (articles dans des revues, sur HAL, etc.), d'un site de valorisation etc.
- Décrire les fichiers de données et chaque variable présente, sa source par exemple, sa modification dans le temps, les données manquantes, etc. (ce que ne permet pas Dataverse par exemple).



## D'autres avantages à diffuser les données historiques par PROGEDO

Centralisation des données sur un portail non institutionnel dont le catalogue compte près de 2000 références.

Diffusion des données sous condition au sein de la communauté scientifique:

l'utilisateur signe un formulaire dans lequel il s'engage à citer l'auteur des données, ainsi qu'à utiliser les données à des fins

La possibilité de diffuser des données pseudonymes : création des fichiers pseudonymisés pour les données qui sont à l'origine non diffusables car soumises au RGPD.

Profiter de l'accompagnement des ingénieurs des PUD et de l'équipe d'Adisp pour la préparation des données et de la documentation de diffusion.



# Comment déposer ses données

progedo-adisp.fr/depot\_donnees.php

PROGEDO - Production et gestion des données en sciences sociales

FR EN



Contact



Aperçu des données

Commander des données

Accueil

ADISP

Données disponibles

Produits à façon

Accès aux données

Dépôt de données

Données internationales

## Comment déposer des données ?

Dans le cadre de l'ouverture des données de la recherche, Quetelet-Progedo Diffusion propose aux chercheurs et aux institutions disposant de données en sciences humaines et sociales de les diffuser à la communauté scientifique. L'équipe Adisp assure la présentation et la diffusion des données selon les standards européens et en respectant les principes FAIR.

Quetelet-Progedo Diffusion a comme vocation d'accueillir tant les données anonymisées que pseudonymisées nécessitant une gestion fine d'accès.

Le dépôt de données est encadré par une convention qui engage les utilisateurs à respecter la citation, l'anonymat et la diffusion des données.

Quetelet-Progedo Diffusion n'est pas propriétaire des données qu'il diffuse et qui restent acquises aux personnes ou institutions ayant déposé les données.

Pour déposer vos données ou toute question, contactez l'Adisp : [diffusion.adisp@cnr.fr](mailto:diffusion.adisp@cnr.fr)

Le dépôt des fichiers doit être accompagné d'une documentation suffisante pour en permettre la réutilisation :

- La description de l'enquête et des fichiers de données (notice ci-jointe),
- Le dictionnaire des variable contenant les noms des variables, leurs labels et leurs descriptions,
- Tout document écrit susceptible d'apporter de l'information supplémentaire sur la démarche de recherche.

Vous pouvez consulter quelques conseils en vue du dépôt de vos données.

Les fichiers de données doivent de préférence être aux formats CSV, Excel ou SAS. Un système de partage de fichiers sécurisé est disponible pour la transmission des données.

Dernière mise à jour : 19 novembre 2021



## Les étapes du dépôt des données:

1. Contactez votre PUD puis notre équipe: **diffusion.adisp@cnrs.fr**

Nous vous proposerons un rdv zoom ou en présentiel pour discuter sur votre projet de recherche et les données à diffuser. Un conseil scientifique de PROGEDO doit se prononcer sur votre demande de diffusion.

2. Nous vous demanderons ensuite de préparer les données et une documentation suffisante pour les décrire :

- Consultez un document téléchargeable « Quelques conseils » en vue de dépôt des données pour vérifier le respect des règles de nommage des fichiers de données et des variables;
- Remplissez la notice téléchargeable en ligne qui renseigne sur votre recherche et vos fichiers de données ;
- Etablissez un dictionnaire des variables;
- Fournissez tout autre document susceptible d'apporter de l'information supplémentaire sur votre projet de recherche (liste complète de sources, une présentation du projet de recherche, publications etc)

3. Envoyez vos données par un lien sécurisé (FileSender)

4. Nous procéderons ensuite à la rédaction et la signature d'une convention de diffusion entre l'institution de l'équipe de recherche ou chercheur et le CNRS.

5. Nous intégrerons vos données et métadonnées sur nos sites et portail Quetelet-Progedo Diffusion et les ouvrirons à la diffusion.



# Un exemple de la page d'enquête diffusée sur le portail PROGEDO

progedo-adisp.fr/enquetes/XML/ill.php?ill=ill-1471

PROGEDO - Production et gestion des données en sciences sociales

archivage de données /  
de la statistique publique

Aperçu des données Commander des données

Accueil ADISP Données disponibles Produits à façon Accès aux données Dépôt de données Données internationales

## Mariages parisiens - 1885

La base est construite à partir d'un échantillon de mille actes de mariages enregistrés dans les mairies des vingt arrondissements de Paris et dans quelques villes du département de la Seine au mois de janvier 1885. Les données permettent une analyse des caractéristiques démographiques et socio-professionnelles de la population parisienne de la fin du XIXe siècle.

Données disponibles / ill-1471

### Descriptif de l'enquête

**N° d'identification :** ill-1471  
**Producteur(s) :**  
 • Garden, Maurice ( ) - ENS Cachan

### Citation bibliographique

**Modèle de citation :** Mariages parisiens - 1885 - Garden, Maurice (producteur, auteur), ADISP (diffuseur).

*La source des données diffusées doit apparaître dans les publications s'appuyant sur ces données (notes de bas de pages, références bibliographiques, légendes de tableaux).*

### Étendue de l'enquête

**Classification :**

- Données historiques
- HISTOIRE
- DÉMOGRAPHIE (POPULATION, ÉTAT CIVIL, ET RECENSEMENT)

**Mots-clés :**

- mariage
- Paris
- banlieues
- XIXe siècle
- villes
- familles

**Date(s) de collecte :** fin = 1997

#### Couverture géographique :

- Paris
- Département de la Seine

**Univers :** C'est à l'évidence une population artificielle, construite par nous, et par son passage devant le maire, comme on dit familièrement, dans un espace donné ; les vingt arrondissements de Paris et dix-neuf communes du département de la Seine, et dans un temps limité (quoique légèrement inégal d'un lieu à l'autre), c'est-à-dire le mois de janvier 1885. Trois «sous-populations» principales sont dégagées par la mise en fiche systématique de mille actes de mariage. La première, la seule vraie ment complète, est celle des «jeunes» mariés : elle est composée de 2 000 personnes, 1 000 hommes et 1 000 femmes que leur histoire individuelle réunit ce jour-là. La seconde est constituée par la parenté directe, les ascendants, père set mères des mariés. Le troisième groupe est constitué par les témoins, obligatoirement 4 par mariage.

**Unité géographique :** arrondissements, commune.

**Unité d'analyse :** Individu.

**Type de données :** Données administratives individuelles

**Nombre de fichiers :** 3

#### Méthodologie et traitement

**Dimension temporelle :** Janvier 1885

**Organisme responsable de la collecte des données :**

- Garden, Maurice

**Fréquence de collecte :** quotidienne

**Méthode d'échantillonnage :** Nous avons retenu un petit échantillon de mariages parisiens de l'année 1885, en opérant une double péréquation, en fonction de la population des vingt arrondissements de Paris, et des deux arrondissements du département de la Seine (Sceaux et Saint-Denis), et en fonction du nombre de mariages enregistrés dans chaque mairie ; au total notre échantillon est ainsi composé de 1 000 mariages, tous du mois de janvier 1885, les premiers de chaque registre, jusqu'au dernier numéro exigé par l'application de cette péréquation (de 25 pour les arrondissements les moins peuplés, Ier, IIe, XVIe, jusqu'à 72 pour le plus peuplé, le XIe), en tout 800 mariages de la ville de Paris, et 200 des communes du département de la Seine hors Paris.

**Mode de collecte :** Saisie sur Excel

#### Version de l'enquête

**Version des fichiers :** Version 1

**Date de la version diffusée :** 2021-11-23

#### Documents relatifs à l'enquête

- ✓ Dictionnaire des variables base témoins
- ✓ Dictionnaire des variables base mari
- ✓ Dictionnaire des variables base femme

Tous les documents disponibles seront communiqués avec le(s) fichier(s) de données.

#### Accès à la documentation des variables sur Nesstar

Le logiciel Nesstar permet d'accéder à l'information au niveau de chaque variable (intitulé de la question, modalités et fréquence des réponses, univers, instructions aux enquêteurs, etc.).

# Un aperçu des données sur le site de Nesstar

← → ↻ Non sécurisé | nesstar.progedo-adisp.fr/webview/index.jsp?object=http://nesstar.progedo-adisp.fr:80/obj/fStudy/lil-1471

**disp** **QUETELET PROGEDO** DIFFUSION

ADISP Commander des données Toutes les enquêtes

DESCRIPTION TABLEAU ANALYSE

Mariages parisiens - 1885

- ☐ Métadonnées
  - ☐ Description de l'étude
    - \* Citation bibliographique
    - \* **Étendue de l'étude**
    - \* Méthodologie et processus
    - \* Accès aux données
  - ☐ Autre matériel pour la description de l'étude
    - ☐ Publications reliées
      - \* Garden Maurice. Mariages parisiens à la fin du XIXe siècle : une micro-analyse quantitative. In: Annales de démographie historique, 1998-1. Le mariage, règles et pratiques, pp. 111-133.
  - ☐ Description des fichiers de données
    - \* GFEMM
    - \* GTEM
    - \* GMARI
  - ☐ Autre documentation
  - ☐ Description des variables

Jeu de données: Mariages parisiens - 1885

---

**Résumé**

La base est construite à partir d'un échantillon de mille actes de mariages enregistrés dans les mairies des vingt arrondissements de Paris et dans quelques villes du département de la Seine au mois de janvier 1885. Les données permettent une analyse des caractéristiques démographiques et socioprofessionnelles de la population parisienne de la fin du XIXe siècle.

---

**Liste des mot-clés**

- mariage
- Paris
- banlieues
- XIXe siècle
- villes
- familles

---

**Classification des sujets**

- Données historiques
- HISTOIRE
- DÉMOGRAPHIE (POPULATION, ÉTAT CIVIL, ET RECENSEMENT)

---

**Date de collection**

Début	Fin	Cycle
	1997	

---

**Pays**

France

---

**Étendue géographique**

- Paris
- Département de la Seine

---

**Unité géographique**

arrondissements, commune.

---

**Unité d'analyse**

Individu.

---

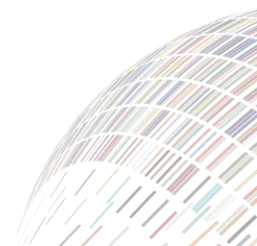
**Univers**

C'est à l'évidence une population artificielle, construite par nous, et par son passage devant le maire, comme on dit familièrement, dans un espace donné ; les vingt arrondissements de Paris et dix-neuf communes du département de la Seine, et dans un temps limité (quoique légèrement inégal d'un lieu à l'autre), c'est-à-dire le mois de janvier 1885. Trois « sous-populations » principales sont dégagées par la mise en fiche systématique de mille actes de mariage. La première, la seule vraie ment complète, est celle des « jeunes » mariés : elle est composée de 2 000 personnes, 1 000 hommes et 1 000 femmes que leur histoire individuelle réunit ce jour-là. La seconde est constituée par la parenté directe, les ascendants, père set mères des mariés. Le troisième groupe est constitué par les témoins, obligatoirement 4 par mariage.

---

**Type de données**

Données administratives individuelles





# Des statistiques simples consultables.

**ADISP** Commander des données

**QUETELET  
PROGEDO  
DIFFUSION**

Mariages parisiens - 1885

- Métadonnées
  - Description de l'étude
    - Citation bibliographique
    - Étendue de l'étude
    - Méthodologie et processus
    - Accès aux données
  - Autre matériel pour la description de l'étude
  - Description des fichiers de données
  - Autre documentation
  - Description des variables**
    - Base mari
      - Numéro d'ordre
      - Arrondissement de mariage
      - Date de mariage
      - Mois de mariage
      - Jour de mariage
      - Nom du mari
      - Prénom 1 du mari
      - Prénom 2 du mari
      - Prénom 3 du mari
      - Prénom 4 du mari
      - Prénom 5 du mari
      - Date de naissance
      - Année de naissance
      - Mois de naissance
      - Jour de naissance
      - Lieu de naissance de mari
      - Code de département du lieu de naissance de mari
      - Age du mari
      - Médaille 1 du mari
      - Médaille 2 du mari
      - MARIPROFES
      - Code profession mari
      - Adresse du mari: numéro
      - Adresse du mari: numéro bis
      - Adresse du mari: voie
      - Adresse du mari: nom de la voie
      - Adresse du mari: Ville
      - Adresse du mari: Code département
      - Adresse précédent du mari
      - Prénom 1 du père
      - Prénom 2 du père
      - Prénom 3 du père
      - Prénom 4 du père
      - Prénom 5 du père
      - Age du père
      - Profession du père**
        - Code profession
        - Médaille 1 du père
        - Médaille 2 du père
        - Médaille 3 du père
        - Père décédé
        - Lieu de décès du père

DESCRIPTION	TABLEAU	ANALYSE
PERCEUR	1	0,2%
PERRUQUIER	1	0,2%
PHOTOGRAPHE	1	0,2%
PLOMBIER	1	0,2%
POLISSEUR SUR CUIVRE	1	0,2%
PORCELAINIER	1	0,2%
PORTIER	1	0,2%
PROFESSEUR	2	0,3%
PROFESSEUR DE CHANT	1	0,2%
PROPRIETAIRE	44	6,7%
PROPRIETAIRE CHATEAU DE CAILA (33)	1	0,2%
RECEVEUR DE L'OCTROI	1	0,2%
RENTIER	26	4,0%
REPRESENTANT DE COMMERCE	2	0,3%
RESTAURATEUR	1	0,2%
RETRAITE	2	0,3%
RETRAITE DES DOUANES	1	0,2%
SCIEUR DE LONG	3	0,5%
SCIEUR DE PIERRES	1	0,2%
SECRETAIRE	1	0,2%
SELLIER	2	0,3%
SELLIER HARNACHEUR	1	0,2%
SERRURIER	6	0,9%
SERRURIER EN COFFRE-FORT	1	0,2%
SP	19	2,9%
SURVEILLANT EN CHEF MAISON DE SANTE DE CLERMONT	1	0,2%
TAILLEUR	11	1,7%
TAILLEUR DE CRISTAUX	1	0,2%
TAILLEUR DE LIMES	1	0,2%
TAILLEUR DE PIERRE	1	0,2%
TAILLEUR DE PIERRES	2	0,3%
TANNEUR	1	0,2%
TAPISSIER	1	0,2%
TARAUDEUR	2	0,3%
TEINTURIER EN BONNETERIE	1	0,2%
TERRASSIER	5	0,8%
TISSERAND	7	1,1%
TISSEUR	2	0,3%
TONNELIER	1	0,2%
TOURNEUR	2	0,3%
TOURNEUR SUR BOIS	1	0,2%
TRAITEUR	1	0,2%
TYPOGRAPHE	1	0,2%
VANNIER	1	0,2%
VERNISSEUR	1	0,2%
VERRIER	2	0,3%
VIGNERON	7	1,1%
VOYAGEUR DE COMMERCE	1	0,2%

**STATISTIQUES SOMMAIRES**

Cas valides	652
Cas manquants	0

Cette variable est alphanumérique

Merci pour votre attention.

