



HAL
open science

Quetelet-Progedo diffusion : Trouver et accéder aux données

Elodie Pétorin

► **To cite this version:**

Elodie Pétorin. Quetelet-Progedo diffusion : Trouver et accéder aux données. Semaine Data-SHS, Dec 2021, Lyon, France. hal-03476894

HAL Id: hal-03476894

<https://hal.science/hal-03476894v1>

Submitted on 3 Jan 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License



DATA
INFRASTRUCTURE

*Développer
la culture des données*

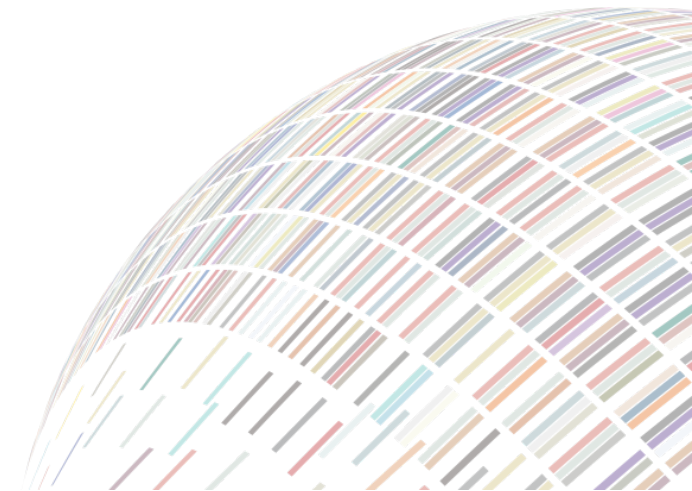


Quetelet-Progedo diffusion : Trouver et accéder aux données

Semaine DATA-SHS, PUD-PANELS

Lundi 6 décembre 2021

Elodie Pétorin, ADISP-PROGEDO



Déroulé de la présentation

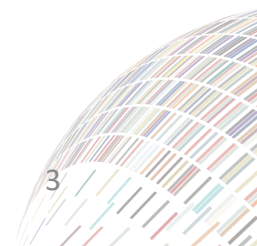
- I. Le **périmètre** de Quetelet-Progedo Diffusion
 - Préambule : public visé, type de données*
 - Les 4 partenaires*
 - Les données françaises*
 - Les données internationales*

- II. Les **informations disponibles** sur les données de Quetelet
 - Les sites web*
 - Les sites Nesstar*
 - Les moteurs de recherche*

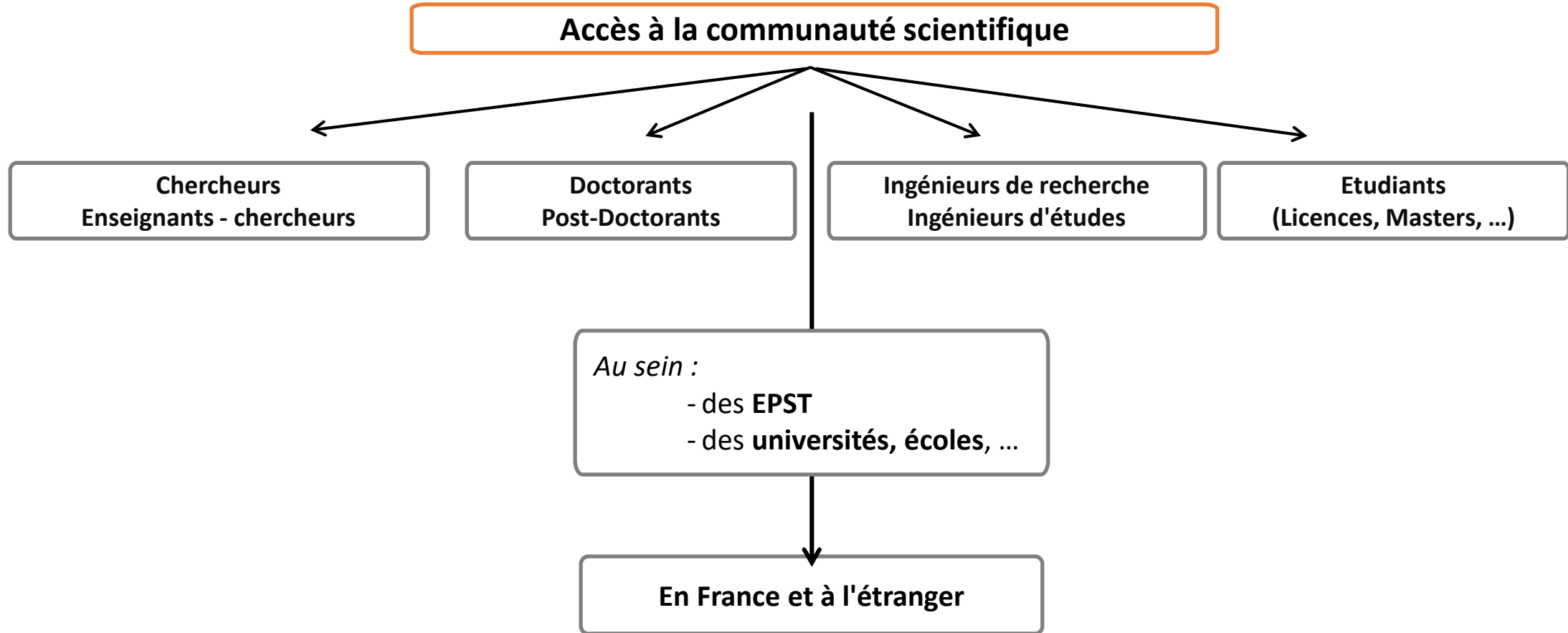
- III. La **commande** de données
 - La nouvelle procédure "FPR"*
 - L'application de commande Quetelet*
 - L'accès via le CASD*

- IV. *Pour en savoir plus ...*

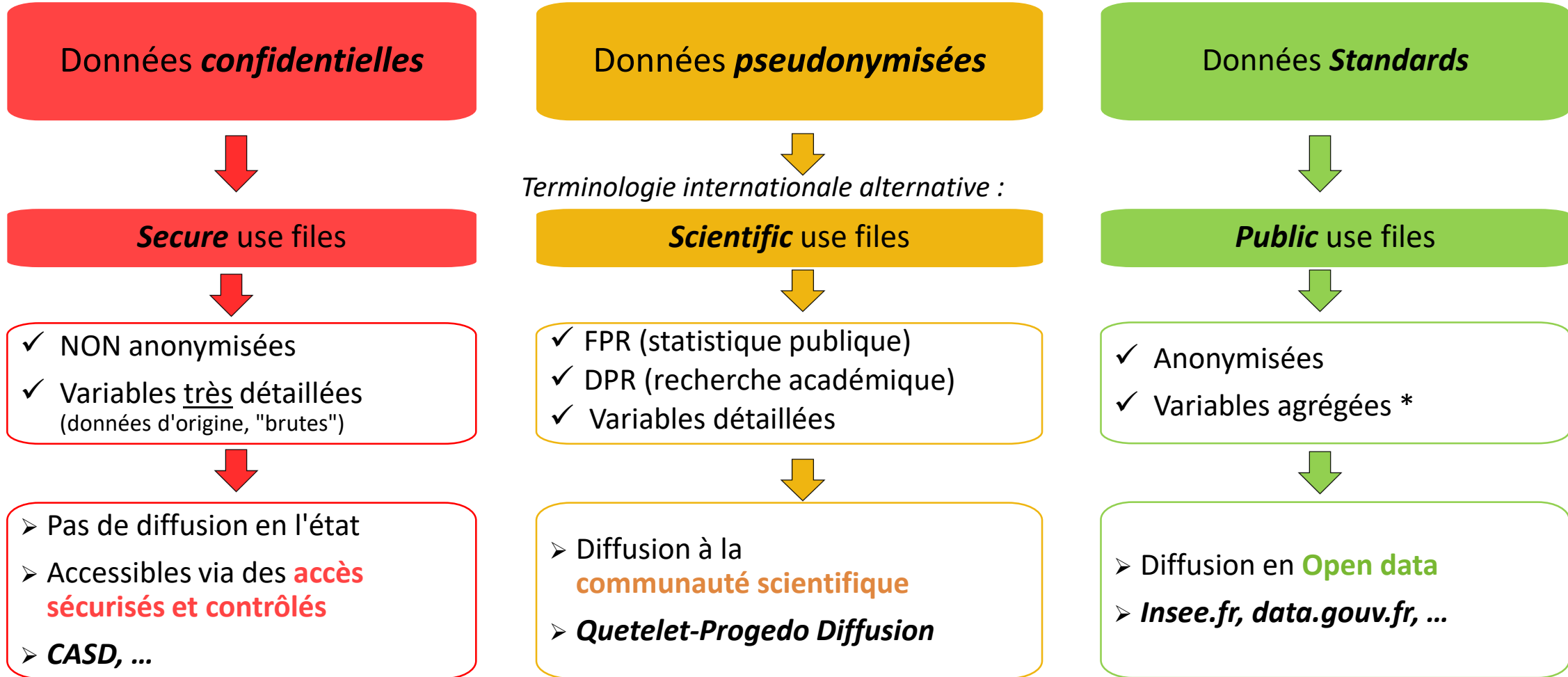
I. Le périmètre de Quetelet-Progedo diffusion



Les publics "autorisés"

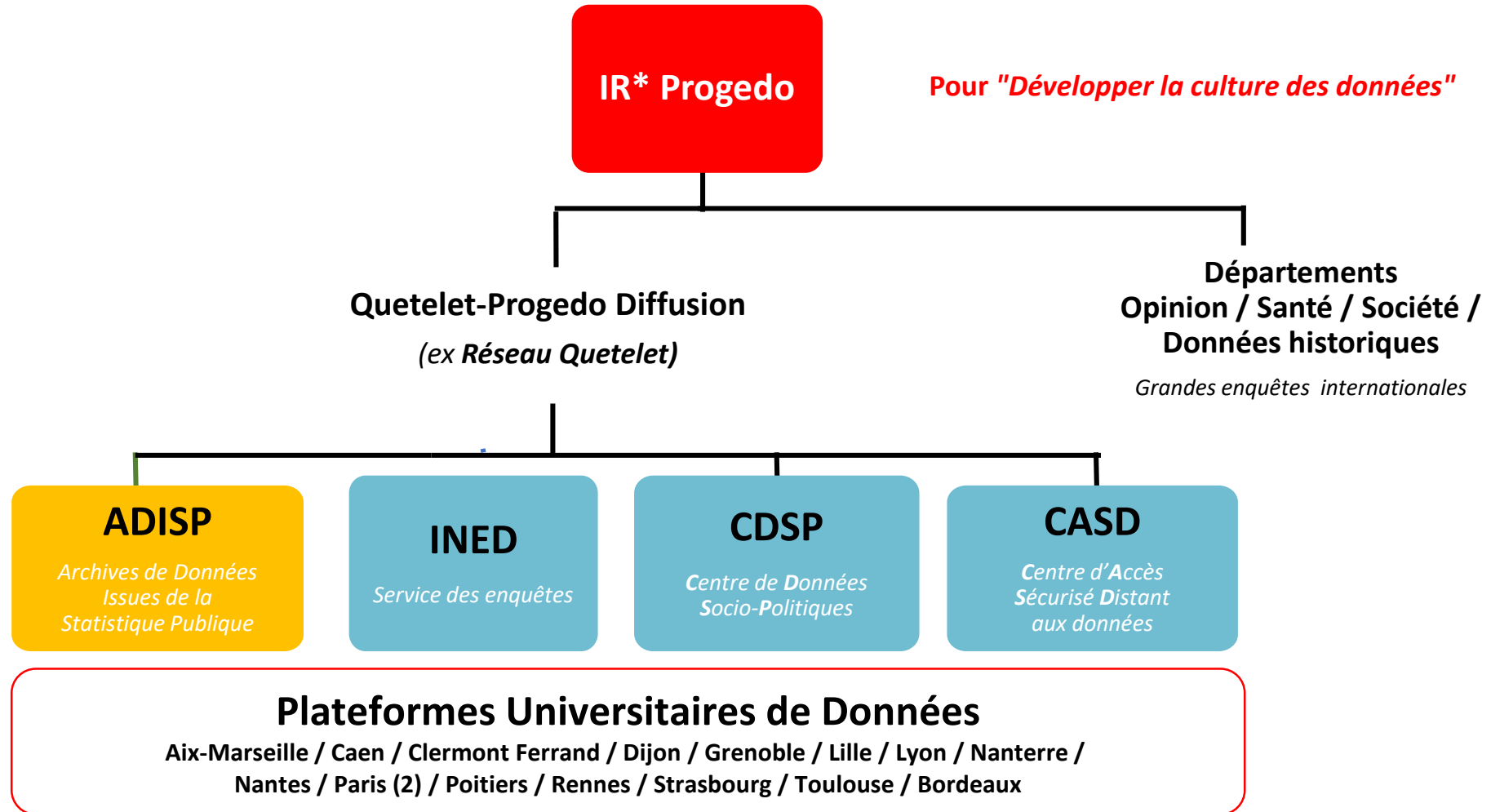


Les types de fichiers



(*) Agrégations le plus souvent opérées sur les variables géographiques, de professions et autres nomenclatures ; ainsi que sur des variables numériques synthétisées en tranches

Introduction : le périmètre de Quetelet-Progedo Diffusion



Catalogue du CDSP

Près de **350 références** disponibles
.....
2 types de **données** :

Données **quantitatives**



- Résultats électoraux (142) :**
- Présidentielles (*depuis 1965*)
 - Législatives (*depuis 1958*)
 - Municipales (*depuis 1983*)
 - Régionales (*depuis 1986*)
 - Référendums (*depuis 1992*)
 - ...
- Enquêtes quantitatives (190)**
- Sur 21 thèmes proposés
 - De 1958 à 2020
 - Champ : France
- *Dt enquêtes via **Elipss***

Données **qualitatives**



Via 



- 17 enquêtes disponibles
- D'autres en préparation ...



Catalogue de l'Ined

INED
(service des enquêtes)



Producteur / diffuseur de ses enquêtes
.....
65 références disponibles *
autour de 8 thématiques

** fin 2020*



- Couples, famille, sexualité
- Naissance, fécondité, contraception
- Espérance de vie, mortalité, santé
- Migrations, discrimination, intégration
- Habitat, environnement, territoire
- Genre, inégalités hommes-femmes
- Ages de la vie, vieillissement
- Populations du monde

Catalogue de l'ADISP : les producteurs



DEPS
Ministère de la Culture

DSED
interieur.gouv.fr
MINISTÈRE DE L'INTÉRIEUR



MEOS
MINISTÈRE
DES SPORTS



DEPP
MINISTÈRE
DE L'ÉDUCATION
NATIONALE

SIES
Ministère de
l'Enseignement Supérieur

SOes
Ministère de la Transition
écologique et solidaire



CEPREMAP
CENTRE POUR LA RECHERCHE ÉCONOMIQUE ET SES APPLICATIONS



Catalogue de l'ADISP : les types de données

environ 220 sources



Près de 1500 références
(Insee : 75% / SSM : 15% / Autres : 10%)



- **80 à 100 nouvelles références / an**
- **et 30 à 50 mises à jour / an**

Uniquement des données quantitatives

✓ **enquêtes et/ou bases administratives**

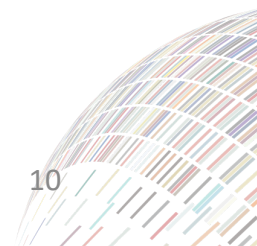
✓ **uniques ou répétées en série**

✓ **de 1954 à 2020**
(et la population en 1850 - 1900)

✓ **métropole et/ou DOM**

✓ **données individuelles ou agrégées**

✓ **certaines, dans un cadre européen**



Catalogue de l'ADISP : les 10 thèmes

Population et démographie



RP / Etat civil / Trajectoires et origines / ...

Travail et emploi



Enq. Emploi / Conditions de travail / Génération / ...

Salaires et revenus



DADS / Structure des salaires / Revenus fiscaux et sociaux / ...

Education et formation



FQP / Panel du 2nd degré / Base centrale de scolarité / ...

Conditions de vie et société



Enq. Logement / Budget des familles / Emploi du temps / ...

Santé et protection sociale



Santé et soins médicaux / Santé des élèves / Handicap / ...

Opinions



Baromètre DREES / ECAM / ISSP / ...

Entreprises



Démographie des entreprises / RECME / ...

Données localisées



EGT / EMD / Impôt sur le revenu dans les communes... / ...

Données historiques



Pop. et enseignement : la France au XIX^e / Base "TRA" / ...

Autres produits via l'ADISP

Projections de population (Omphale)

- **projections démographiques** à moyen / long terme
(de population, de ménages ou de population active)
- sur **tout territoire supra-communal** de plus de 50 000 habitants
- à partir d'hypothèses sur la fécondité, la mortalité et les migrations

Distancier METRIC

- Calcul de **distances** et de **temps de parcours** ...
... d'une commune à une autre (données supracommunales)
... d'un point à un autre (avec données géolocalisées)

Produits Sur Mesure (PSM)

- tableaux à **la demande**
- à partir d'une vingtaine de sources de l'INSEE
(recensement, état civil, revenus, emploi, entreprises, etc.)
- fichiers non disponibles dans le catalogue de l'ADISP
- incluant des **informations plus détaillées** que l'offre standard

Autres produits via l'ADISP : le "sur mesure"

Les PSM = Produits Sur Mesure

Tableaux à la demande sur fichiers ...
... non disponibles dans le catalogue de l'ADISP
... **plus détaillés** que les offres standard / FPR

à partir de 19 sources INSEE

Recensement de la population

- Recensement de la population de 2006 à 2014
- Historique des recensements de la population (Saphir) et

État civil

- Naissances
- Décès
- Mariages
- Reconnaissances

Données sur les entreprises

- Stocks d'entreprises et d'établissements (REE/Sirene)
- Créations d'entreprises et d'établissements (REE/Sirene)
- Transferts d'établissements (REE/Sirene)
- Système d'information sur les nouvelles entreprises (Sine)
- Enquête sur les liaisons financières entre sociétés (Lifi)
- Système unifié de statistiques d'entreprises (Suse)

Fonction publique

- Système d'information sur les agents des services publics (Siasp)
- Recensement des agents de l'État
- Enquête sur les personnels des collectivités territoriales et des établissements publics locaux (Colter)

Travail, emploi, salaires

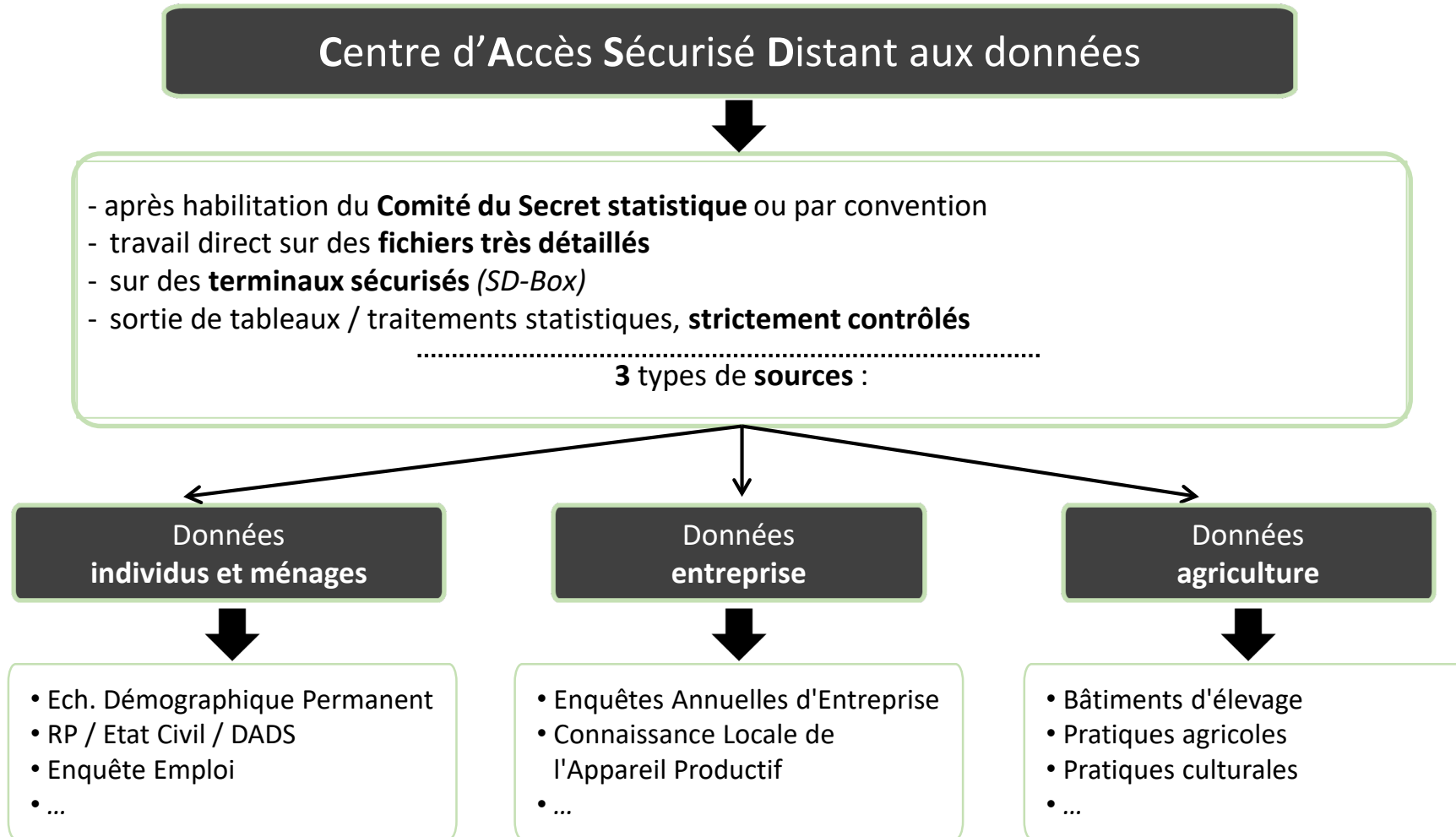
- Déclaration annuelle de données sociales (DADS)
- Connaissance locale de l'appareil productif (Clap)
- Urssaf - Extension du Projet Urssaf pour les revenus et l'emploi (Epure)

Revenus

- Enquête Revenus fiscaux et sociaux (ERFS)

(Consulter les années disponibles sur le [site de l'ADISP](#))

Le CASD : service rendu et catalogue



II. Les informations disponibles sur les données de Quetelet

Les sites "institutionnels" des 4 partenaires



Focus sur : *Données de la statistique publique*



Focus sur :

- *Données quantitatives*
- *Résultats électoraux*
- *Ressources en ligne*



Focus sur : *La recherche Enquêtes Ined*



• *Données qualitatives*



Focus sur : *Données « brutes »*

Consultation des données : via "Nesstar"



- En bref :
- environ 2/3 des références présentées (avec détail des variables)
 - pas de possibilité de tris à plat, croisés, ...



- En bref :
- Enquêtes quantitatives du CDSP, avec détail des variables
 - possibilité de traitements statistiques



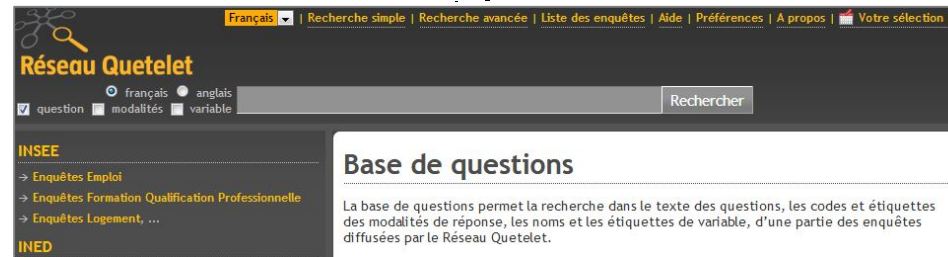
- En bref :
- totalité du catalogue de l'INED, avec ou sans le détail des variables
 - description générale des enquêtes non mises à disposition
 - possibilité de traitements statistiques

Les moteurs de recherche



Plan du site

| QUI SOMMES NOUS ? | RESSOURCES DISPONIBLES | ACCÉDER AUX DONNÉES | DÉPOSER SES DONNÉES |
|---|---|---|--|
| <ul style="list-style-type: none"> LE CONTEXTE LES 4 DIFFUSEURS LES MISSIONS | <ul style="list-style-type: none"> DONNÉES FRANÇAISES DONNÉES INTERNATIONALES | <ul style="list-style-type: none"> RECHERCHER DES DONNÉES CONDITIONS D'ACCÈS COMMANDER DES DONNÉES | <ul style="list-style-type: none"> À QUI S'ADRESSER DOCUMENTER POUR LA RÉUTILISATION CONSERVER À LONG TERME |



Précisions :

- ✓ recherche sur la **description générale**
- ✓ **toutes** les références **ADISP et INED**
- ✓ **données quantitatives** du **CDSP**

Précisions :

- ✓ Requête sur les **questions, variables, modalités**
- ✓ **uniquement** les références **entièrement documentées** (*i.e. jusqu'au niveau des variables*)

Exemple de recherche dans la BdQ (1/2)

Requête : "cinéma"

157 résultats, dont :

| | | | |
|--|---|--|---|
| <p>EPCV 2006 de 1996 à 2004 :</p> | <p><i>Combien de fois en moyenne êtes-vous allé au cinéma au cours des douze derniers mois ? (5 500 à 6 000 individus par an)</i></p> | <p><i>(nombre de fois par jour, semaine ou mois)</i></p> | <p>15 ans ou plus en ménages ordinaires, en France métropolitaine</p> |
| <p>SRCV 2006 - 2009 :</p> | <p><i>Au cours des douze derniers mois..., êtes-vous allé(e) au cinéma ? Il s'agit ici de sorties au cinéma et non pas de visionnage de films à la télévision, sur DVD... (20 000 ind.)</i></p> | <p>- Oui - Non</p> | <p>16 ans ou plus en ménages ordinaires, en France métropolitaine</p> |
| <p>Emploi du temps :</p> | <p>1974 : Etes-vous <i>déjà</i> allé au cinéma ? (6 600)</p> | <p>- Oui / - Non</p> | <p>18 ans ou plus en communes urbaines</p> |
| | <p>1985 : Etes-vous <i>déjà</i> allé au cinéma ? (10 000)</p> | <p>- Oui / - Non</p> | <p>15 ans ou plus en France métropolitaine</p> |
| | <p>1998 : Au cours des 4 dernières semaines, êtes-vous allé au cinéma ? (16 000)</p> | <p>- Oui / - Non</p> | <p>15 ans ou plus en France métropolitaine</p> |
| | <p>2009 : Au cours des 4 dernières semaines, êtes-vous allé au cinéma ? (18 000)</p> | <p>- Oui / - Non</p> | <p>11 ans ou plus ; métropole et DOM</p> |

Exemple de recherche dans la BdQ (2/2)

| | | | |
|---|--|--|--|
| Handicap - Santé 2008 : | <i>Au cours des douze derniers mois, êtes-vous allé au cinéma ? (30 000 individus)</i> | <ul style="list-style-type: none"> - Oui - Non - Ne sait pas | De tous âges en situation de handicap et France entière |
| Conditions de vie des étudiants de 1994 à 2006 : | <i>Au cours des 30 derniers jours, êtes-vous allé au cinéma ? (25 000 individus)</i> | <ul style="list-style-type: none"> - Oui - Non | Inscrits dans les étab. d'ens. sup. français ; métropole et DOM |
| Gai Pied Hebdo 1990 : | <i>Vous sortez au moins au cinéma ... (2 000 individus)</i> | <ul style="list-style-type: none"> - 1 fois par semaine - 1 fois tous les 15 jours - 1 fois par mois - moins souvent | Homo- et bisexuels , en France |
| Violence envers les femmes 2000 : | <i>Au cours du dernier mois, êtes-vous sortie seule le soir pour aller au cinéma ? (7 000 individus)</i> | <ul style="list-style-type: none"> - Oui - Non | Femmes de 20 à 59 ans, en France métropolitaine |

Limites actuelles de l'exercice :

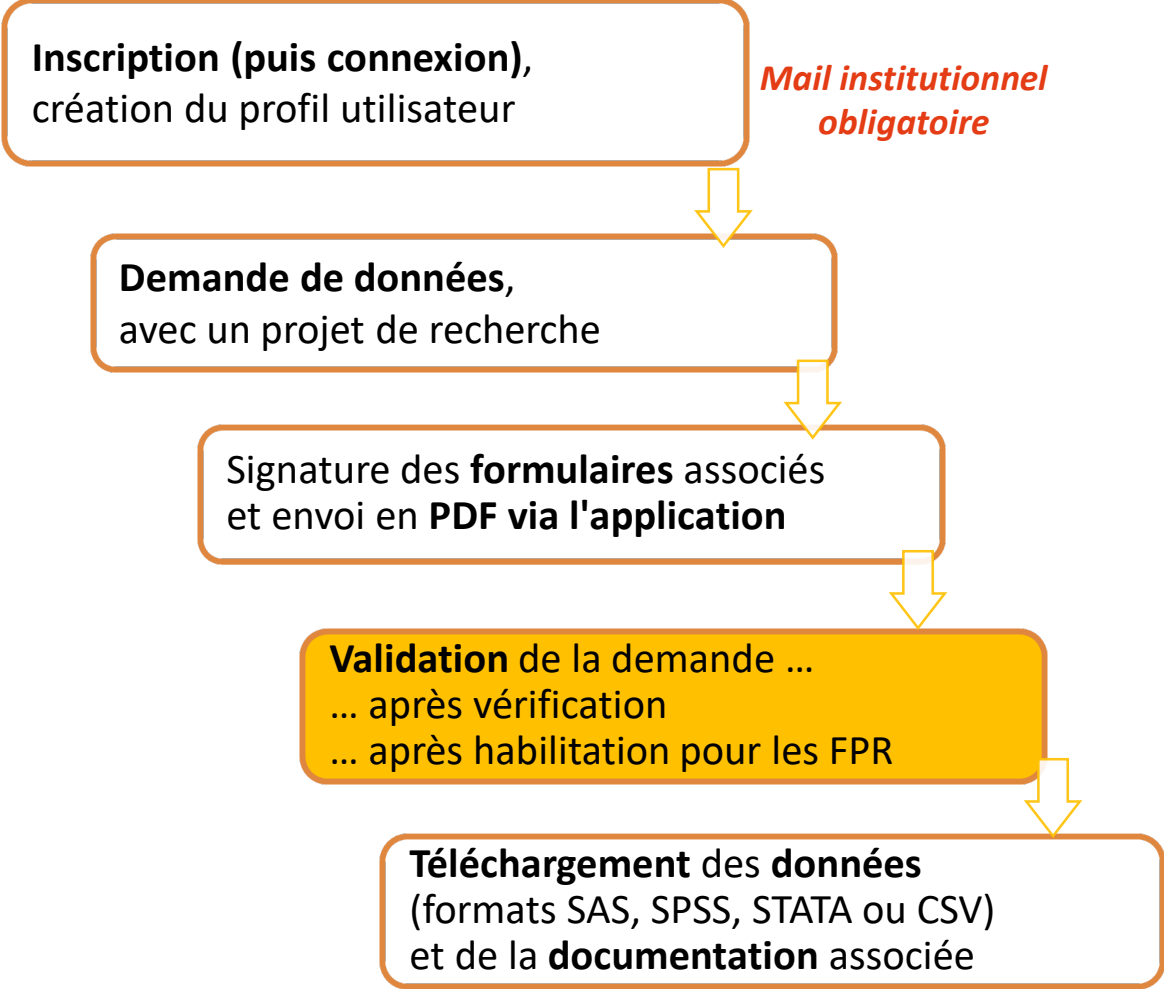


Pas de mention de l'enquête sur les Pratiques culturelles ...

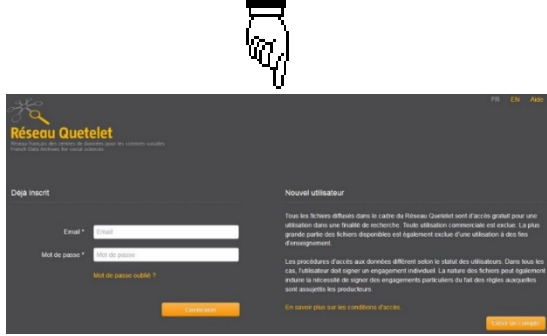
III. La commande de données

L'application de commande unifiée (hors CASD)

Etapes des demandes de données via l'application de commande de Quetelet (pour les données diffusées par **l'ADISP, le CDSP et l'INED**)

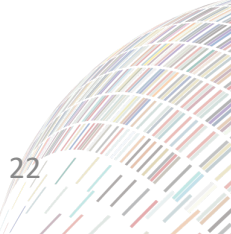


*Mail institutionnel
obligatoire*

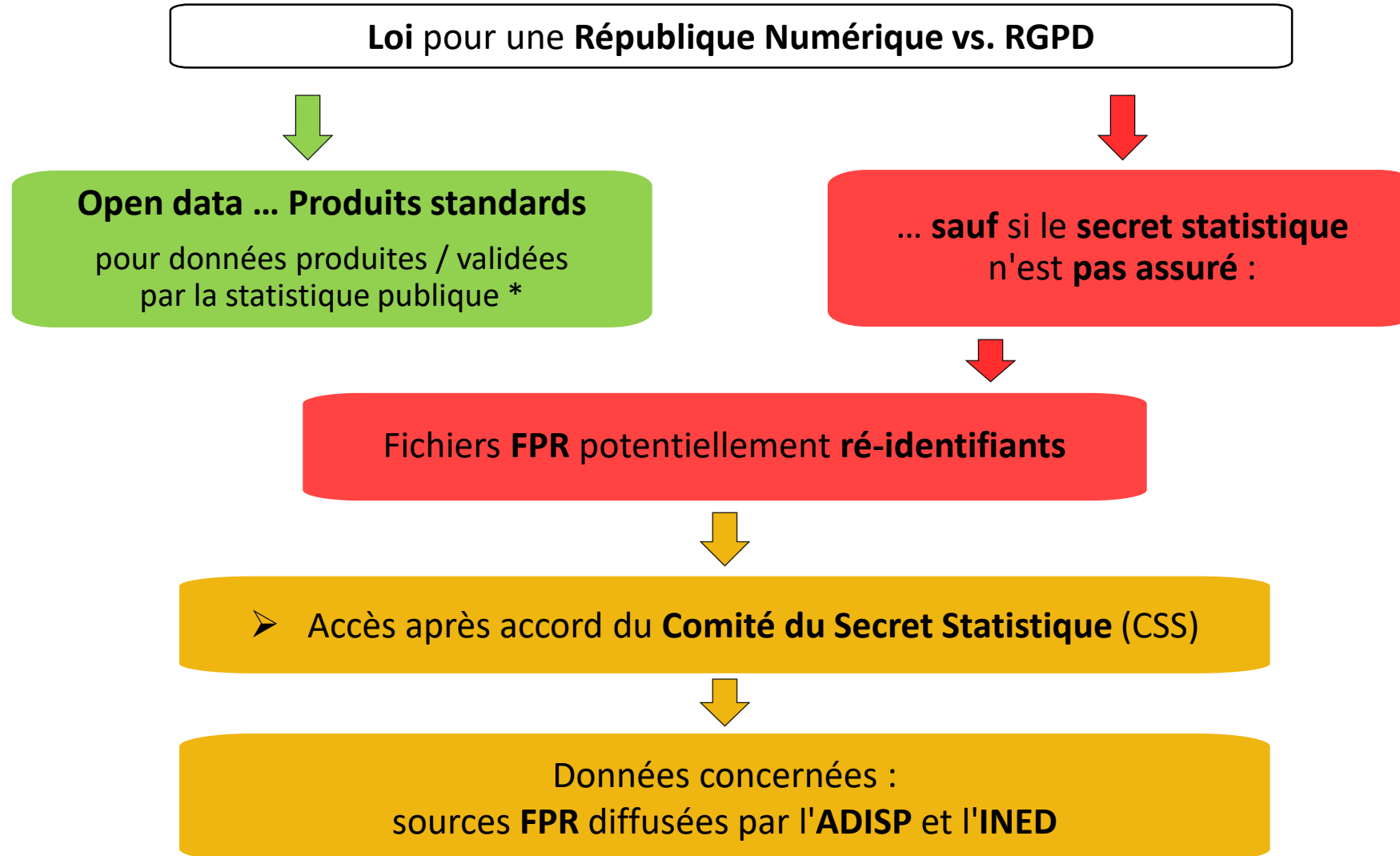


Coût : gratuit,
y compris PSM,
Omphale, Metric

Délai :
- 1^{ère} demande FPR :
2 à 3 semaines
- Autres demandes FPR
et données standards :
2 à 3 jours



Spécificité pour la diffusion des "FPR"



(*) statistique publique = passage par le CNIS et/ou le comité du label

Pourquoi une habilitation ?

Toutes les données non anonymes nécessitent une habilitation du comité du secret statistique (CSS), donc mise en place d'une **procédure simplifiée** et spécifique pour l'accès aux FPR.

Concrètement

Nous sommes **l'unique interlocuteur** pour le chercheur, l'étudiant.

Coté utilisateur : signer le formulaire d'engagement spécifique du CSS

Coté diffuseur : vérifier que le demandeur appartient à un organisme de l'annexe 2 (établissement de l'ESR + établissement étranger), sinon nous demandons à ajouter l'établissement.

Envoi des demandes au CSS qui doit valider et faire signer aux archives.

Dès qu'on a le retour (environ 2-3 semaines), on valide les demandes.

L'habilitation reste valable pour l'ensemble des FPR tant que le demandeur appartient au champ de l'annexe 2.

Données confidentielles : process via le CASD

Les étapes d'une demande d'accès via le CASD

Identification des sources de données nécessaires, description du projet
+ **contact avec le(s) service(s) producteur(s)**



Soumission du dossier de demande complet envoyé au secrétariat du comité, via le **portail CDAP**



Étude du dossier par le comité du secret et avis
(lors de séances plénières ou de consultations électroniques)

+ Signature de l'accord par le service
producteur et les Archives



Obtention d'une carte d'accès lors d'une séance **d'enrôlement** et contractualisation auprès du CASD



Accès aux données : **Traitement** des **données** via la SD-box



Validation des traitements par le CASD



Accès aux tris, graphiques, sorties

Une procédure pour faciliter l'utilisation des données par les étudiants

L'enseignant peut anticiper leur demande d'habilitation

en fournissant la liste des intéressés (*nom – prénom, mail universitaire*), afin que leur procédure d'habilitation soit initiée rapidement (par exemple en juin ou juillet pour un cours prévu en septembre ou octobre) sur notre mail diffusion.adisp@cns.fr

Cette anticipation ne remplacera pas les **demandes que les étudiants devront ensuite faire individuellement** (même dans le cadre d'un travail de groupe) sur le portail de commande, mais celles-ci pourront alors être traitées plus vite (2 à 3 jours après finalisation de la demande).

Données "FPR" : Obligations de l'utilisateur

Dans la demande de FPR, l'utilisateur s'engage à :

- ✓ **Utiliser les données exclusivement dans une finalité de recherche**
- ✓ **Ne pas céder ces données**, sous quelque forme que ce soit, à une tierce personne, que ce soit à titre gratuit ou onéreux
- ✓ Traiter ces données conformément aux **règles de l'art et du secret statistique**
- ✓ **Mentionner la source des données** dans les communications, publications ... conformément au modèle de citation pour l'utilisation des données
- ✓ Informer le diffuseur des communications, publications ... et lui en faire parvenir un exemplaire et les références
- ✓ Informer le diffuseur des constats relatifs à la qualité des données ou à leur difficulté d'utilisation
- ✓ Informer le diffuseur de toute réutilisation des données pour une autre recherche que celle spécifiée ci-dessus
- ✓ **Détruire les fichiers à l'issue du travail de recherche**
- ✓ **Respecter la réglementation en matière de protection des données personnelles.**

**Merci de votre
attention**

Elodie Pétorin

elodie.petorin@cnrs.fr ou

diffusion.adisp@cnrs.fr

