



HAL
open science

L'expérience de l'entrepôt Nesstar au Centre de données socio-politiques (CDSP)

Alina Danciu, Alexandre Mairot

► To cite this version:

Alina Danciu, Alexandre Mairot. L'expérience de l'entrepôt Nesstar au Centre de données socio-politiques (CDSP). Journée nationale de la Science Ouverte, Atelier Fairisation des entrepôts de données 19 novembre 2019, Nov 2019, Paris, France. hal-03901299

HAL Id: hal-03901299

<https://hal-sciencespo.archives-ouvertes.fr/hal-03901299>

Submitted on 15 Dec 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.

L'EXPÉRIENCE DE L'ENTREPÔT NESSTAR AU CDSP (Sciences Po, PROGEDO)

Journée nationale de la Science Ouverte,
Atelier Fairisation des entrepôts de données
19 novembre 2019

Alina Danciu, Alexandre Mairot

Les entrepôts NESSTAR

Qu'est-ce que Nesstar ?

Nesstar est un logiciel de publication de données et d'analyse en ligne. Le logiciel se compose d'outils qui permettent aux fournisseurs de données de diffuser leurs données sur le Web. Nesstar gère les données d'enquête et les tableaux multidimensionnels ainsi que les ressources textuelles. Les utilisateurs peuvent rechercher, parcourir et analyser les données en ligne.

Que fait Nesstar ?

- Publier des données et des métadonnées
- Donner accès à toutes les données par le biais d'un système unique
- Permettre aux utilisateurs d'analyser ou de télécharger des données en ligne
- Visualisez les données à l'aide de cartes, graphiques et tableaux

Source : <http://www.nesstar.com/about/about.html>

L'entrepôt NESSTAR au CDSP



SciencesPo
CENTRE DE DONNÉES SOCIO-POLITIQUES

CDSP

- ELIPSS
- French Electoral Study
- Démocratie 2000
- Les militants socialistes à l'épreuve du pouvoir (1984-1986)
- Regional Income Inequality in the United States (1917-2011)
- Enquête post-électorale de l'élection présidentielle 2012
- Baromètre Politique Français 2006-2007
- Panel électoral français 2002
- Enquêtes électorales françaises
- Enquêtes interrégionales des phénomènes politiques
- Enquêtes OIP Jeunes et Seniors
- Enquêtes auprès des élus
- General election study Belgium
- Enquêtes Image de la science
- Enquêtes IFEN sur l'environnement
- Presse Gaie
- L'ouvrier français en 1970
- Database regarding the North Indian MPs since 1952
- Enquêtes auprès des policiers
- The European Union Legislative Output (1996-2014)
- Enquêtes Agoramétrie (1977-2005)
- Sondages de Sortie des Urnes 1995

Catalogue du Centre de données socio-politiques

Ce site permet un accès en ligne, direct et gratuit, à des enquêtes d'utilisation fréquente, documentées conformément à la [norme internationale DDI](#) (Data Documentation Initiative) et organisées grâce au logiciel NESSTAR.

Aucune inscription n'est nécessaire pour la consultation de la documentation des données et des tris à plat, tris croisés ou analyses simples. Si vous souhaitez télécharger un jeu de données,

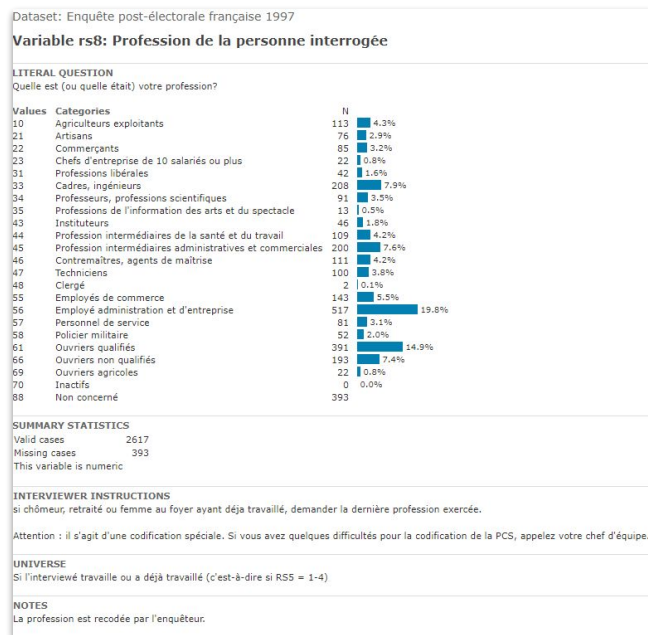
[inscrivez-vous sur le portail des données du réseau Quetelet dont le CDSP fait partie.](#)

Vous pouvez accéder ci-dessous à [une documentation d'utilisation de Nesstar](#) :



Source : <http://nesstar.sciences-po.fr/webview/>

La description des données



Pour chaque variable :

- Le libellé de la question exacte ou la signification exacte de la donnée
- Le texte de la question
- Information univers c'est à dire qui a répondu à la question
- Signification exacte des codes
- Codes des données manquantes
- Statistiques de distribution de fréquences non pondérées
- Imputation et édition de l'information (connaître les informations qui ont été modifiées)
- Détails sur les constructions et les variables de poids
- Situation dans le fichier de données
- Groupement des variables

La provenance des données

Une liste non exhaustive de la **description du contexte de production** des données dans l'entrepôt :

- Personnes responsables de la création des données
- Un résumé décrivant le but, la nature et la portée des données
- Termes clés qui décrivent des aspects importants des données
- Les grands sujets importants couverts par les données
- Renseignements sur les subventions
- L'organisme désigné pour produire des copies des données
- Période à laquelle se réfèrent les données
- Tout matériel lié à l'ensemble des données
- Couverture géographique des données
- Unité d'analyse ou d'observation de base
- Description de la population couverte par les données du fichier
- Type d'échantillon et plan d'échantillonnage
- Type d'instrument de collecte
- Pourcentage des membres de l'échantillon qui ont fourni des renseignements

Les conditions de réutilisation

Availability Status

Les données sont disponibles depuis septembre 2019 sur le portail Quetelet-Progedo Diffusion.

Confidentiality Declaration

L'obtention des données est soumise à la signature d'un engagement qui prévoit, entre autres :

- le respect de la confidentialité des répondants ;
- la non-rediffusion des données à un tiers ;
- la citation de la source des données dans les publications ;
- la destruction du fichier de données à l'issue de la recherche.

Restrictions

Les données ne peuvent être utilisées qu'à des fins de recherche et d'enseignement. Toute autre utilisation, commerciale ou non, est strictement interdite. La finalité s'apprécie par le descriptif du projet de recherche ou d'enseignement qui accompagne chaque demande de données. Les résultats obtenus à partir des données du panel ELIPSS sont publics et libres de diffusion, la finalité ne sert pas les intérêts particuliers d'institutions publiques ou privées.

Access Authority

Name	Affiliation	E-mail address	Universal Resource Identifier
Centre de Données Socio-Politiques	FNSP/CDSP	info.cdsp@sciencespo.fr	http://cdsp.sciences-po.fr/

Deposit Requirement

Les utilisateurs doivent obligatoirement signaler les publications réalisées à partir du panel ELIPSS au Centre de Données Socio-Politiques de Sciences Po en envoyant la référence bibliographique et le résumé.

Conditions

Les procédures d'accès diffèrent selon le statut du demandeur. Dans tous les cas, l'acceptation de la demande faite depuis le portail du réseau Quetelet conduit à la signature d'un engagement individuel.

Disclaimer

Le Centre de Données Socio-Politiques de Sciences Po, qui produit et diffuse les enquêtes du panel ELIPSS, n'est pas responsable des interprétations ou déductions faites à partir de l'utilisation des données.

L'interopérabilité avec



- La *Data Documentation Initiative*, projet international né en 1995, développe et maintient un standard pour la description et la conservation des données.
- Les spécifications DDI, écrite en XML, fournit un format pour le contenu, l'échange et la conservation des informations liées à l'étude, ses résultats et les données elle-même.
- Le standard est lui-même interopérable avec d'autres : SDMX, Dublin-Core...
- 3 versions : DDI 2 ou DDI *Codebook*, DDI 3 ou DDI *Lifecycle* et le futur DDI 4

Les vocabulaires contrôlés du cessda

Actuellement 7 des 10 vocabulaires en cours d'intégration dans la description et la provenance des données. Ils concernent :

- l'unités d'analyse
- le format général des données
- le mode de collecte
- la procédure d'échantillonnage
- la méthode temporelle
- la thématique des données
- le type d'instrument