



DATA
INFRASTRUCTURE

PROGEDO

Rapport d'activité 2021

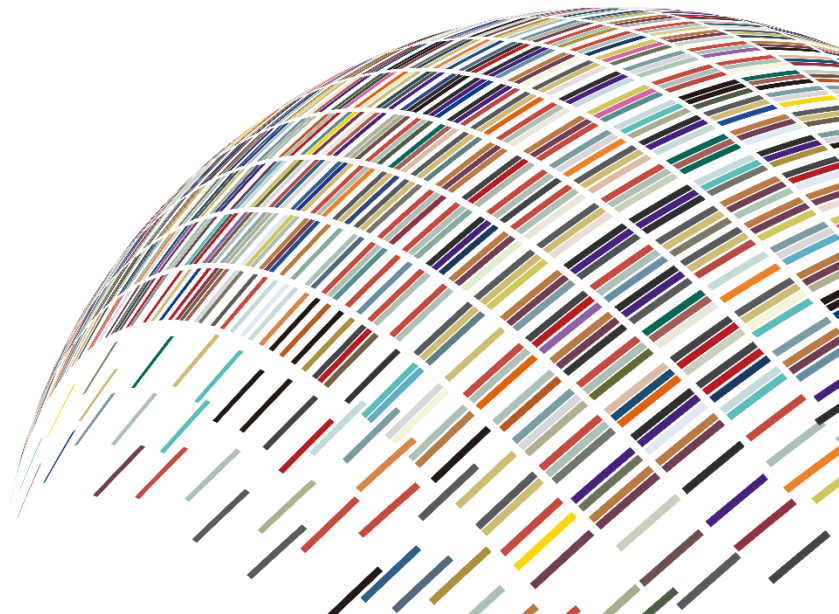


Table des matières

INTRODUCTION	2
MISSION ET ORGANISATION DE L'IR*	3
MISSION	3
ORGANISATION	3
ORGANIGRAMME	4
PRODUCTIONS DES DONNEES	6
LES GRANDES ENQUETES INTERNATIONALES.....	6
<i>ESS</i>	6
<i>SHARE</i>	7
<i>ISSP</i>	8
<i>GGP</i>	9
LES DONNEES FRANÇAISES	10
<i>Elipss</i>	10
<i>D-FIH</i>	11
<i>Popp</i>	11
GESTION ET PARTAGE DES DONNEES	13
QUETELET-PROGEDO-DIFFUSION	13
<i>Archives des Données Issues de la Statistique Publique</i>	14
<i>CASD</i>	15
AUTRES ACCES AUX DONNEES	16
<i>INSEE</i>	16
<i>ICPSR</i>	16
<i>LIS</i>	16
PROMOTION DES DONNEES	18
COMMUNICATION DE L'IR* PROGEDO	18
<i>Sites web</i>	18
<i>Carnet hypothèse</i>	18
<i>canal U</i>	19
<i>HAL</i>	19
PLATEFORMES UNIVERSITAIRES DE DONNEES.....	20
<i>Accompagnements individuels</i>	21
<i>Accompagnements de projets</i>	21
<i>Formations</i>	21
<i>Séminaires et conférences</i>	21
<i>Semaine Data SHS</i>	21
<i>Ateliers de la donnée</i>	22
LES DEPARTEMENTS.....	23
<i>Méthodes et statistiques publiques</i>	23
<i>Opinions et enquêtes et sociopolitiques</i>	23
<i>Santé</i>	24
<i>Données historiques économiques et financières</i>	24
<i>Population et parcours de vie</i>	24
LIFE OBS	25
BILAN FINANCIER	26
RESSOURCES	26
DEPENSES	27

Introduction

Après une année 2020 compliquée par la pandémie et les fermetures des universités qui ont perduré jusqu'à l'été 2021, l'IR* PROGEDO a accentué son travail de structuration du champ des données issues d'enquêtes et de bases de données quantitatives en sciences sociales. Très concrètement, les enquêtes en face-à-face ont pu reprendre, plusieurs projets ont abouti (par exemple celui de la refonte du catalogue de données de Quetelet-Progedo-Diffusion pour correspondre aux standards du CESSDA), de nouvelles plateformes de données ont démarré. Les actions ont concerné aussi bien l'unité (USR 2006) qui sert d'appui à l'infrastructure de recherche, que son réseau de plateformes universitaires de données et ses départements d'animation. Ce rapport d'activité rend compte du travail accompli, qu'il a fallu résumer en quelques pages, qui rendent compte trop brièvement de l'ensemble du travail fourni partout en France et au niveau européen.

L'ensemble de ce travail s'est effectué en synergie avec les autres infrastructures de recherche en SHS. L'IR RnMSH, qui accueille les plateformes universitaires de données au sein de ses maisons, est le premier partenaire de l'IR* PROGEDO. La relation avec l'IR* HumaNum est vitale pour PROGEDO, puisqu'elle accueille et gère nos serveurs et que les outils proposés à l'ensemble de la communauté SHS sont nos outils de travail quotidiens. Comme pour toute unité de l'InSHS, l'IR OpenEdition constitue un soutien important, notamment à travers les « carnets hypothèses ». A leur côté PROGEDO participe au projet Opera (<https://www.operas-eu.org/>).

La qualité du travail fourni par l'ensemble des personnes impliquées dans l'IR* PROGEDO que ce soit au niveau de la production des enquêtes, de leur documentation, de leur diffusion ou de leur promotion a permis à PROGEDO d'être à nouveau confirmée sur la feuille de route nationale des infrastructures de recherche pour la 5^{ème} fois (après 2008, 2012, 2016 et 2018). Un travail de remise à plat de la gouvernance de l'IR* a été mené avec l'appui de son comité de pilotage, pour correspondre au modèle des autres Infrastructures de recherche française. 2022 verra l'avènement de cette démarche avec un changement de l'unité d'appui, désormais unité propre du CNRS (UAR 2506).

Rendez-vous est donc pris pour 2022.

Sébastien Oliveau,

Directeur de l'IR* PROGEDO

Mission et organisation de l'IR*

Mission

PROGEDO (pour PROduction et Gestion des DONnées) est une infrastructure de recherche étoile (IR*) visant à développer la culture des données en sciences humaines et sociales. Son champ est centré sur les données d'enquête. Elle soutient la création et la diffusion de grandes enquêtes internationales et l'usage des statistiques publiques par les chercheurs.

Organisation

PROGEDO repose sur une organisation multiscalaire : internationale, nationale, locale.

À l'échelle internationale et européenne, PROGEDO soutient la participation de la France dans la production de grandes enquêtes pluridisciplinaires, notamment les enquêtes ESS (European Social Survey), SHARE (Survey on Health, Ageing and Retirement in Europe), GGP (Generations and Gender Programme), EVS (European Value Study), ISSP (International Social Survey Programme). PROGEDO coordonne en outre la participation française au CESSDA (Consortium of European Social Science Data Archives). Ce consortium qui met en réseau différentes archives de données européennes pour la recherche en SHS, stimule la coopération entre les membres afin de proposer des services utiles au développement de recherches nationales et internationales.

A l'échelle nationale, L'IR* PROGEDO s'organise thématiquement autour de grands départements chargés d'animer des communautés de recherche transverses aux disciplines. Ces départements, animés par un·e enseignant·e-chercheur·e ou un·e chercheur·e, travaillent en synergie entre eux et avec les différents niveaux de la l'IR* : équipes des enquêtes internationales, équipe de la diffusion nationale (QPD), équipes locales des PUD. Chaque département a des objectifs spécifiques qui correspondent aux enjeux de leur thématique en regard des données existantes ou à créer. Ils collaborent avec des équipes partenaires de l'IR* et au-delà des SHS.

À l'échelle française, PROGEDO contribue, par le biais du portail Quételet-PROGEDO Diffusion à répondre aux problématiques d'archivage, de documentation et de diffusion des données françaises issues d'enquêtes ou de la statistique publique pour la recherche en sciences humaines et sociales. Le portail héberge les données issues des Archives de Données Issues de la Statistique Publique (1500 enquêtes), du Centre de Données Socio-Politiques de Sciences Po et du service d'enquête de l'Ined. PROGEDO soutient en outre le Centre d'Accès Sécurisé aux Données qui permet l'accès dans des conditions de sécurité élevées à des données individuelles très détaillées.

Dans les universités, au plus près des jeunes chercheurs, des équipes et des écoles doctorales, l'action de PROGEDO est menée par les Plateformes Universitaires de Données (PUD). Développées avec le Réseau national des Maisons des Sciences de l'Homme, elles s'y développent dans un cadre optimal de décloisonnement et de mutualisation. Les PUD sont réparties sur l'ensemble du territoire métropolitain. Il s'agit d'un dispositif original qui vise à stimuler l'utilisation des données mises à disposition par PROGEDO et qui est le plus à même de développer la culture des données en France en gagnant de nouveaux publics.

Organigramme

Sébastien Oliveau, Directeur

Claire Vandemoortele, Secrétaire Générale

Sandra Brée, en charge du département « données historiques, économiques et financières »

Frédéric Gonthier, en charge du département « Opinion et enquêtes sociopolitiques »

Muriel Roger, en charge du département « Méthodes et statistiques publiques »

Philippe Terral, en charge du département « Santé »

Archives des Données Issues de la Statistique Publique (ADISP) :

Erik Zolotoukhine (responsable de l'équipe),

Lorraine Adam

Elodie Petorin

Alena Lapatniova

Ana Prada Aranguren

Référent de l'enquête ESS-France : Agnaly Michaud (CDSP), coordinatrice nationale

Référent de l'enquête SHARE-France : Florence Jusot (Université Paris-Dauphine), Country Team Leader

Référent de l'enquête GGP-France : Laurent Toulemon (INED), coordinateur national

Responsable des équipes partenaires :

Centre d'Accès Sécurisé aux Données (CASD) : Kamel Gadouche

Centre de Données Socio-Politiques (CDSP) : Nicolas Sauger

Service des enquêtes de l'Institut Nationale des Etudes Démographiques (Ined) : Thomas Merly-Alpa

- Plate-forme Universitaire d'Aix-Marseille (PUD-AMU) : Isabelle Blöss-Widmer
- Plate-forme Universitaire de Données de Caen (PUD-C) : Maxime Marie
- Plate-forme Universitaire de Données de Dijon (PUDD) : Jimmy Lopez
- Plate-forme Universitaire de Données de Grenoble (PUD-GA) : Frédéric Gonthier
- Plate-forme Universitaire de Données de Lille (PUD-L) : Héloïse Petit
- Plate-forme Universitaire de Données de Lyon (Panels) : Sonia Guérin-Hamdi
- Plate-forme Universitaire de Données de Nantes (PROGEDO-Loire) : Jean Rivière

- Plate-forme Universitaire de Données de Nanterre (PUD-N) : Julie Pannetier
- Plate-forme Universitaire de Données de Rennes (PUD-R) : Nolwenn Le Meur & Emmanuelle Leray
- Plate-forme Universitaire de Données de Poitiers (PUD-P) : Liliane Bonnal
- Plate-forme Universitaire de Données de Strasbourg (PUD-S) : Nicolas Cauchi-Duval
- Plate-forme Universitaire de Données de Toulouse (PUD-T) : Philippe terral
- Plate-forme Universitaire de Données de Clermont-Auvergne (PUD-CA) : Simone Bertoli
- Plate-forme Universitaire de Données de Bordeaux (PUD-B) : Christophe Bergouignan

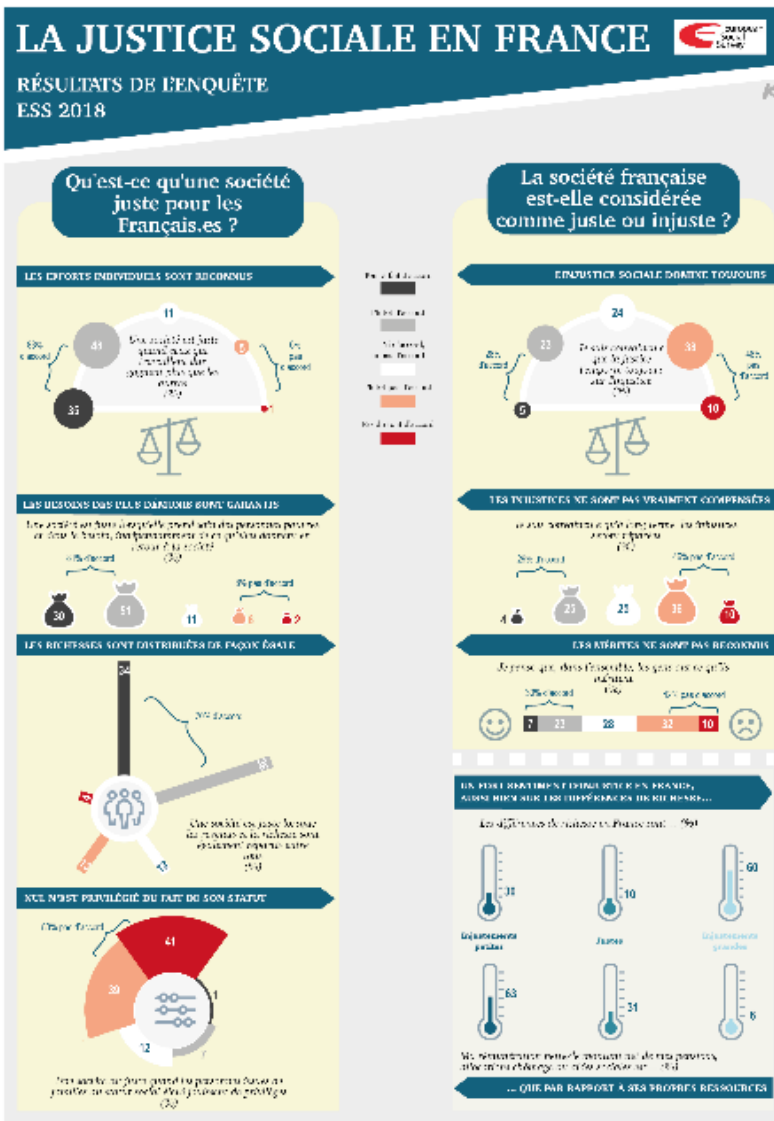
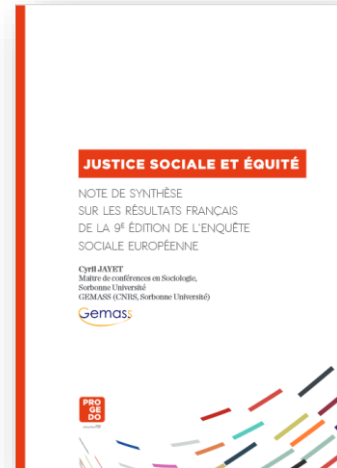
Productions des données

Les grandes enquêtes internationales

ESS

L'European Social Survey (ESS) est un dispositif d'enquête comparative et répétée (depuis 2002) sur les attitudes des Européens. Cette enquête est réalisée à partir de la passation tous les deux ans d'un questionnaire auprès d'échantillons nationaux représentatifs dans près de trente pays. La collecte des données de l'édition 10 de l'enquête a eu lieu entre le mois de septembre et le mois de décembre 2021.

L'équipe de coordination nationale, qui a en charge la mise en œuvre effective de l'enquête sociale européenne sur le terrain, est basée depuis la sixième édition au Centre de données socio-politiques (CDSP) de Sciences Po.



L'enquête sociale européenne (ESS) permet de mieux comprendre le rapport de la société française à la justice sociale et à l'équité grâce à l'enquête sociale européenne. Cyril Jayet (Maître de conférences en Sociologie à Sorbonne Université, membre du GEMASS) a proposé une synthèse très claire grâce dans une note de synthèse rédigée pour PROGEDO (<https://hal.archives-ouvertes.fr/hal-03227402>)

En 2021 PROGEDO a produit une infographie sur le module rotatif consacré à la "Justice Sociale" :

SHARE

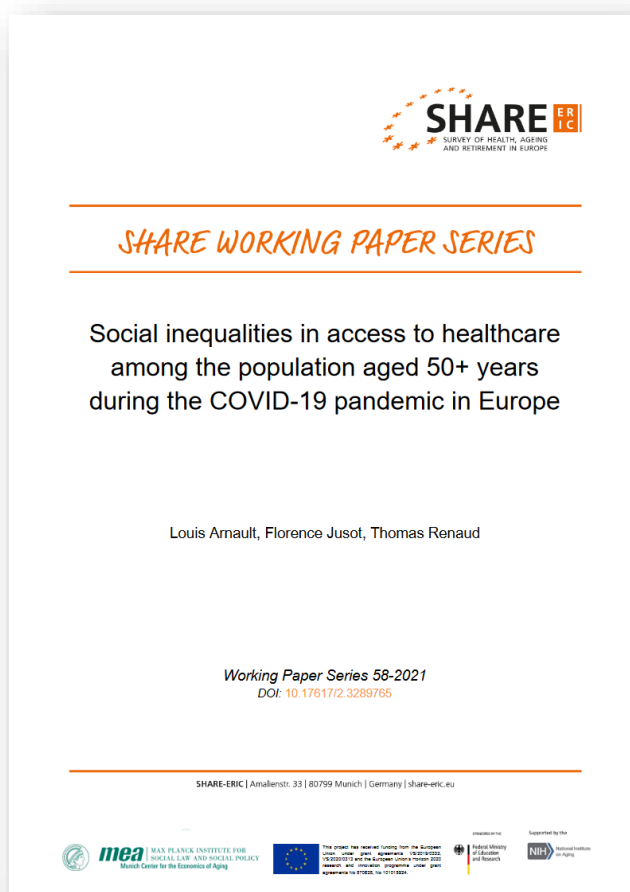
L'enquête européenne SHARE (Survey of Health, Ageing and Retirement in Europe) construit depuis 2004 une base de données longitudinale, multidisciplinaire et internationale concernant plus de 120 000 Européens âgés de 50 ans et plus et ayant pour objectif d'analyser les problèmes économiques et sociaux liés au vieillissement de la population.

Depuis 2012, l'équipe des économistes du LEDa-LEGOS de l'Université Paris-Dauphine a repris la gestion de l'enquête SHARE en France (enquête sur la santé, le vieillissement et la retraite en Europe), jusque-là traitée par l'IRDES (Institut de recherche et de documentation en économie de la santé), et a procédé à la collecte des fonds nécessaires à la poursuite de l'enquête.

Les entretiens de la neuvième vague d'enquête ont commencé en octobre 2021.

Au printemps 2020, une enquête spécifique sur l'épidémie de Covid a été passée par téléphone. Les données sont déjà disponibles et l'enquête a fait l'objet d'un premier article sur les inégalités sociales dans l'accès aux soins pendant la pandémie :

(http://www.share-project.org/uploads/tx_sharepublications/SHARE_WP_58-2021.pdf)



ISSP

L'International Social Survey Programme (ISSP) est un programme d'enquête sociale internationale permettant des comparaisons transnationales et dans le temps. Non thématisée, cette enquête permet de questionner un sujet important pour la recherche en sciences sociales, en complément d'une série de questions sur les caractéristiques sociodémographiques des enquêtés. Elle a lieu annuellement depuis 1996.

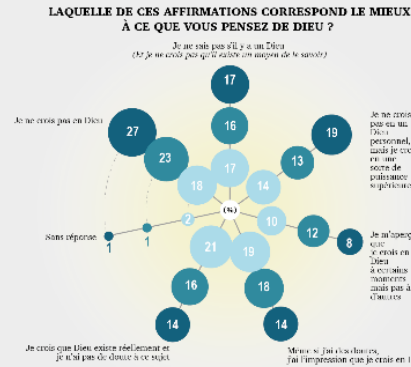
Le volet français de l'enquête est porté en France par une équipe de l'UMR PACTE (CNRS, Université Grenoble Alpes et Sciences Po Grenoble). En 2021, PROGEDO a valorisé le module rotatif sur les religions à travers une infographie :

LA RELIGION EN FRANCE

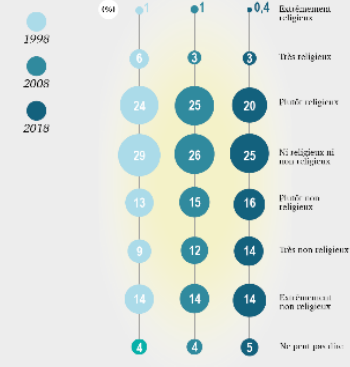


RÉSULTATS DE L'ENQUÊTE
ISSP 2018

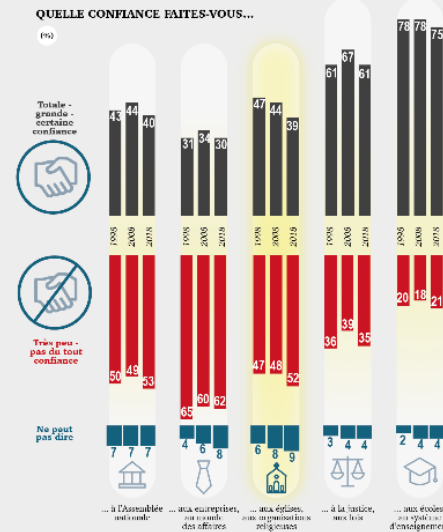
Une croyance en Dieu et une religiosité qui reculent



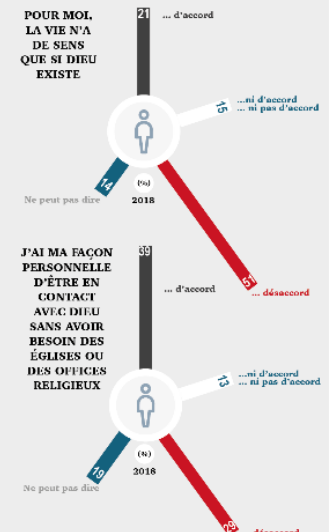
DIRIEZ-VOUS QUE VOUS ÊTES...



Une confiance dans les églises en baisse



Des croyances de plus en plus individualisées



A quoi croit-on ?

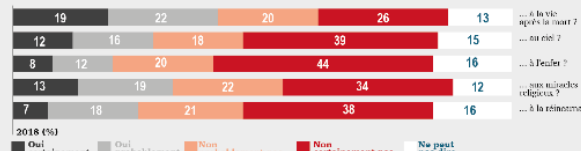
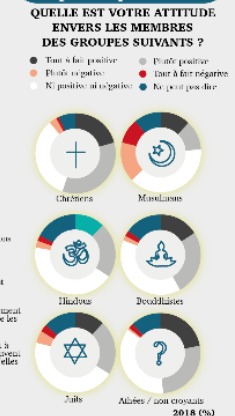
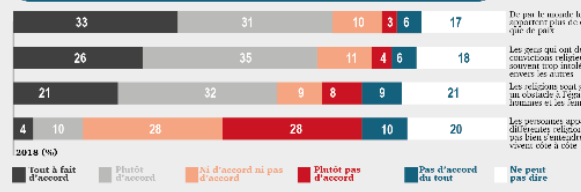


Image des groupes religieux reste plutôt positive



Des religions vues comme source de conflit et d'intolérance



International Social Survey Programme est une enquête internationale réalisée chaque année depuis 1985 sur un thème différent dans près de quarante pays du monde.
Enquête française a été conduite par voie postale entre février et juin 2018 auprès d'un échantillon national représentatif de 1000 ménages. Les résultats sont présentés ici pour les 252 réponses validées reçues. Ils sont redressés en fonction du genre, de l'âge et de la catégorie socio-professionnelle pour compenser les biais d'échantillonnage.
Un rapport détaillé des résultats est disponible sur le site www.progedo.fr.
Les données françaises sont accessibles sur le portail www.insee.fr et les données internationales sur le portail www.issp.ac.uk.
Pacte
Laboratoire de sciences sociales

GGP

Le programme international GGP (Generations and Gender Program) vise à mieux comprendre comment la vie familiale et les relations entre les générations s'organisent différemment, selon la situation économique, l'engagement de l'Etat et les systèmes de valeurs de chaque pays. Ce programme développe l'enquête GGS (générations and gender survey). Le volet français de l'enquête, nommé ERFI (Étude des relations familiales et intergénérationnelles) est portée par l'Institut national d'Études Démographiques (INED). Cette étude s'est adressée aux femmes et aux hommes âgé(e)s de 18 à 79 ans et concerne toutes les situations familiales : en couple ou célibataire, avec ou sans enfants, avec ou sans emploi. La première enquête a eu lieu en 3 vagues entre 2005 et 2011.

Le financement du projet *Life Obs* obtenu dans le cadre du PIA 3 permettra la mise en place d'une nouvelle série 3 de vagues. l'Ined a réalisé en 2021 une enquête pilote auprès de 3.200 personnes représentatives de la population française. Elle a étudié les biais méthodologiques et particulièrement permis une comparaison des modes de passation de l'enquête : CATI (téléphone) vs. CAWI (web). Elle montre la supériorité du téléphone pour limiter les abandons, mais souligne que la robustesse des réponses n'est pas affectée par l'usage d'un mode ou l'autre (résultats présentés par dans le cadre du séminaire GGP-Connect du 25 janvier 2022 par Milan Bouchet-Valat, Marine Bondon, Ruxandra Breda-Popa , Paul Cochet, Géraldine Charrance, Efi Markou et Laurent Toulemon).

Depuis 2021, GGP est inscrit sur la feuille de route nationale des infrastructures françaises en tant que projet ESFRI.

Les données françaises

Elipss

Dans le cadre de l'EquipEx DIME-SHS (<https://dime-shs.sciencespo.fr>), le CDSP a mis en place dès 2012 le panel ELIPSS (Études Longitudinales par Internet Pour les Sciences Sociales), un dispositif d'enquête par internet dédié à la communauté académique. Le panel ELIPSS exploite un échantillon de plus de 2200 personnes (au 1er janvier 2021) qui répondent chaque mois par Internet à des questionnaires conçus par des équipes de recherche en sciences humaines et sociales.

Le projet d'EquipEx a pris fin en janvier 2021, mais le dispositif continue de fonctionner, grâce à la mise en place d'une politique d'autofinancement. PROGEDO intervient comme partenaire depuis sa création à plusieurs titres. En faisant connaître le dispositif aux communautés de recherche d'une part (notamment via ses plateformes universitaires de données), en allant chercher des projets qui pourraient mobiliser cet outil d'autre part (comme c'est le cas pour l'enquête ISSP par exemple), enfin en apportant ponctuellement des soutiens financiers pour aider le dispositif à se maintenir (lors du rafraîchissement du panel par exemple.).

Ce sont désormais plus de 80 enquêtes qui sont disponibles grâce à ce dispositif, et toutes les données sont disponibles : https://cdsp.sciences-po.fr/fr/ressources-en-ligne/?0=subject__id-30

The screenshot shows the 'Ressources en ligne' page on the SciencesPo website. The page features a navigation menu with 'LE CDSP', 'DONNÉES', 'SERVICES', and 'PROJETS'. Below the menu, there is a breadcrumb trail: 'Accueil > Ressources en ligne'. The main content area is titled '82 ressources en ligne' and lists several resources:

- Enquête annuelle ELIPSS - vague 8 (2021)**
Enquête quantitative - 2021
- Faire face au Covid-19 - vague 1 (2020)**
Enquête quantitative - 2020
- Faire face au Covid-19 - vague 2 (2020)**
Enquête quantitative - 2020
- Faire face au Covid-19 - vague 3 (2020)**
Enquête quantitative - 2020
- Faire face au Covid-19 - vague 4 (2020)**
Enquête quantitative - 2020

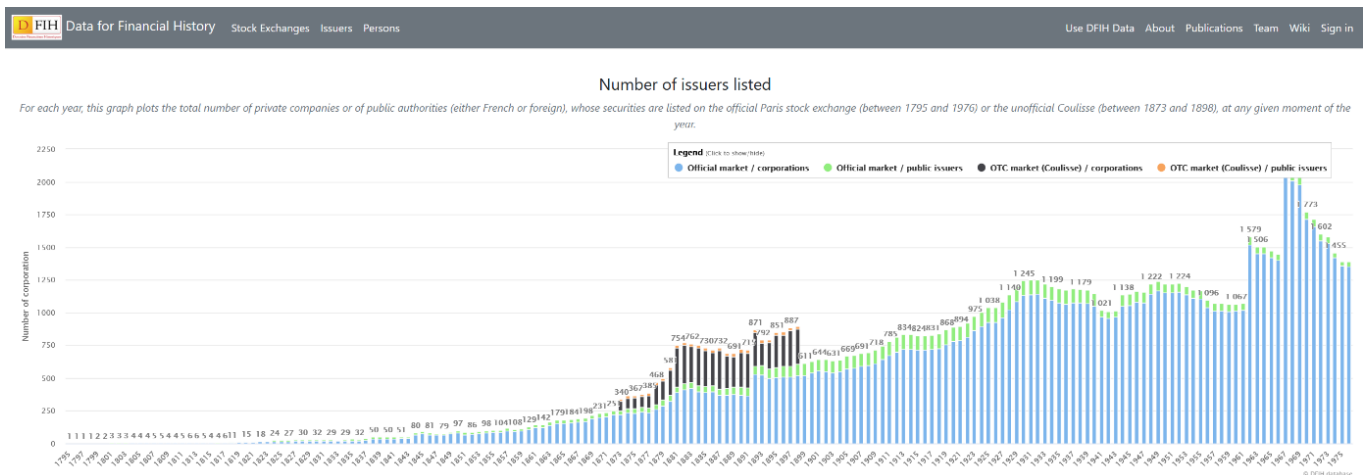
The left sidebar contains filters for 'Type' and 'Thèmes'. Under 'Type', there are three options: 'Enquête qualitative /0', 'Enquête quantitative /82', and 'Résultats électoraux /0'. Under 'Thèmes', there are several options, with 'ELIPSS /82' selected (indicated by a red checkmark). Other themes include 'Administration /4', 'Comportements politiques /79', 'Culture /19', 'Economie /2', 'Éducation /3', 'Elections /144', 'Elus /9', and 'Environnement /21'.

D-FIH

Ce projet vise à sortir d'une situation dans laquelle la quasi-totalité des travaux en finance utilisent la seule base de données sur longue période existante : celle du CRSP à Chicago, qui porte exclusivement sur les actions américaines cotées à New York. Elle n'est donc certainement pas représentative de ce que peut attendre un investisseur ou une entreprise d'un marché financier en général.

Il s'agit de construire une base de données exhaustive sur les titres cotés en France depuis 1800, qui fournisse un point de comparaison par rapport aux données américaines. La base de données DFIH est beaucoup plus riche car elle comporte les actions, obligations, prix au comptant comme à terme, fermes comme optionnels, titres français comme étrangers, pour tous les marchés boursiers (libres ou organisés) français. Elle permet donc d'aborder, outre les questions classiques de la finance, un certain nombre de questions qui restent peu traitées : rôle des bourses régionales, des marchés libres, des interactions entre marchés, choix de cotation des entreprises et des investisseurs, etc.

Financé dans le cadre d'un Equipex (ANR-10-EQPX-0018), La base de données DFIH est librement accessible en ligne (<https://dfih.fr/>) et continue d'être enrichie : de nouvelles données sont ajoutées en permanence. DFIH est à l'origine d'un projet d'infrastructure de recherche européenne : EurHisFirm (<https://eurhisfirm.eu/>).



Popp

Les projets POPP et EXO-POPP visent une meilleure connaissance de la population parisienne et de sa banlieue sous la IIIe République. Ils mettent en place des bases de données spécifiques dont la finalité sera la construction d'un corpus commun.

« Popp » est un premier projet financé par le Collex-Persée (<https://popp.hypotheses.org/>) qui met en place une première base de données de 12 millions d'individus à partir des recensements nominatifs de Paris de 1926, 1931, 1936 et 1946. Ces recensements sont déjà numérisés et en ligne sur le site des archives de Paris mais il s'agit ici de créer une base de données permettant l'exploitation statistique de ces documents particulièrement riches et offrira l'opportunité d'un pas important dans la connaissance de la population urbaine européenne jusqu'alors très peu étudiée.

Un second projet, EXO-POPP (Extraction Optique des entités nommées manuscrites pour les actes de mariage de la population de Paris de 1880 à 1940), a débuté le 1er octobre 2021. Suivre le projet : <https://exopopp.hypotheses.org/>. Il vise à élaborer une vaste base de données à partir de 300 000

actes de mariages de Paris et sa banlieue entre 1880 et 1940. Grâce à la collaboration d'informaticiens et de chercheurs en sciences sociales, ce projet permettra des avancées considérables.

Les derniers progrès en reconnaissance d'écritures manuscrites et en traitement automatique du langage naturel, grâce au Deep Learning, permettent d'envisager des avancées significatives dans la compréhension et l'extraction optique d'entités nommées manuscrites. C'est l'ambition du projet EXO-POPP de développer de nouveaux algorithmes et de les appliquer à l'analyse des actes de mariages qui offrent un cas d'étude exemplaire pour l'utilisation de ces nouvelles techniques.

EXO-POPP va considérablement améliorer la compréhension des transformations familiales sous la III^e République. Au-delà, l'automatisation de la transcription d'archives laisse entrevoir d'immenses possibilités pour la recherche historique quantitative.

DÉSIGNATION		NUMÉROS			NOMS DE FAMILLE	PRÉNOMS	ANNÉE de COMPLÈTE	LIEU de NATURALITÉ	NATIONALITÉ	ÉTAT MATRIMONIAL	SITUATION en chef de ménage	DE BREVÉ D'INSTRUCTION	PROFESSION	Pour les personnes affectées d'incapacité, mentionner à deux colonnes, successivement, le nom du médecin ou du chirurgien qui les constate.
des maisons, cités, villages ou hameaux	des rues, places ou cours	des maisons	des cités	des villages										
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
15 ^e Belleville	24	}	}	}	Colibert	Mathieu	93	Paris	fran	M	ch			1937
						Jeanne	94	Paris	fran	M	ch			1937
						Georget	22	Paris	fran	M	ch			
						Antoinette	27	Paris	fran	M	ch			22405

Recensement de la population de 1926, population de résidence habituelle, arrondissement de Belleville. Cote D2M8 307 (Archives de Paris).

Gestion et partage des données

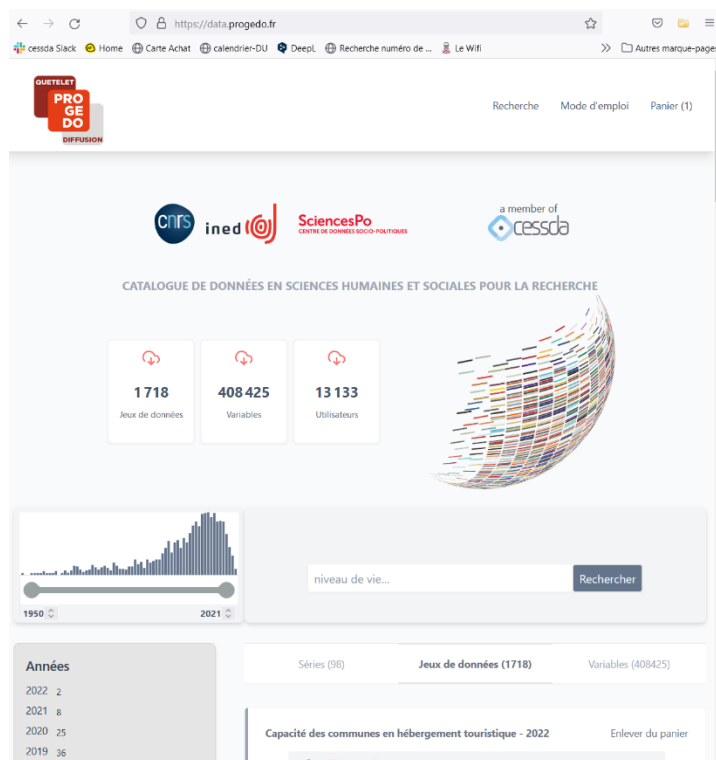
Quetelet-PROGEDO-Diffusion

La diffusion nationale des données pour la recherche est assurée par trois équipes partenaires qui offrent des accès aux grandes bases de données (enquêtes, bases administratives, bases de gestion) provenant de la recherche, de la statistique publique et d'organismes divers:

- les archives des données issues de la statistique publique PROGEDO (CNRS)
- le Service de enquêtes (INED)
- le Centre de Données SocioPolitiques (Sciences Po Paris/CNRS)

Ces équipes coordonnent les activités d'archivage, de documentation et de diffusion des données françaises en sciences humaines et sociales. Quetelet-Progedo-Diffusion participe à l'ERIC CESSDA, qui coordonne les fournisseurs de données européens. Les actions mises en place concourent notamment à :

- la pérennité de l'archivage sur le long terme qui doit faire l'objet d'une veille permanente en lien avec les différents acteurs impliqués sur ces questions dont les Archives nationales,
- l'authentification des données qui doivent faire l'objet de références sans ambiguïté au plan national comme international,
- le recours à des standards professionnels de documentation des jeux de données,
- un meilleur contrôle de la qualité des données, renforcé par les échanges entre producteurs et utilisateurs des données,
- une valorisation des données plus largement utilisées dans des travaux de recherche qui peuvent aussi faire l'objet d'une meilleure validation scientifique du fait de l'accessibilité des données sur lesquelles ces travaux reposent.



Plus de 1700 jeux de données sont actuellement accessibles via Quetelet-Progedo-Diffusion et ses partenaires.

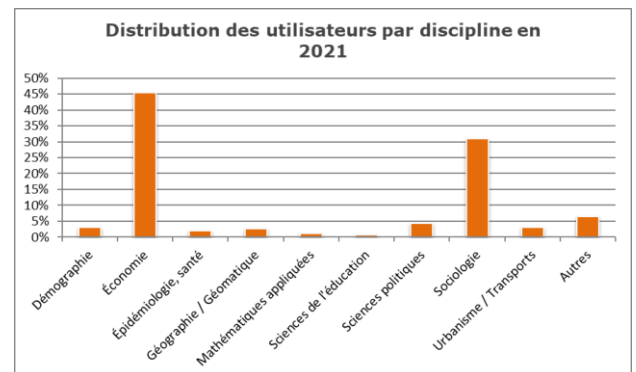
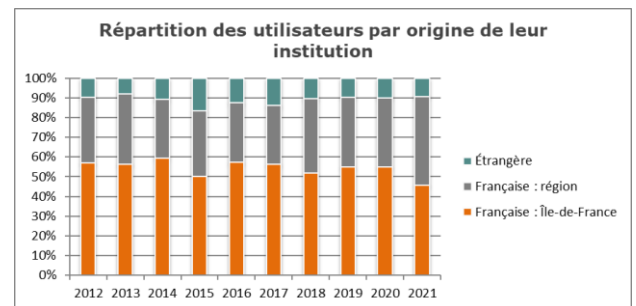
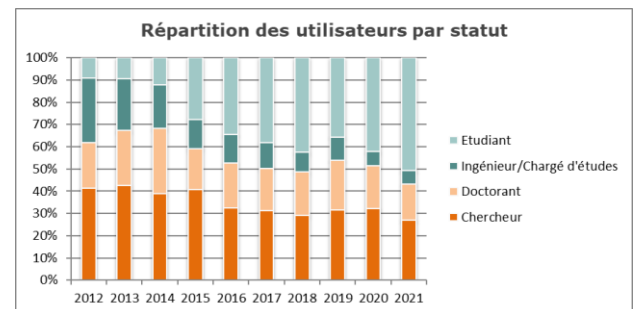
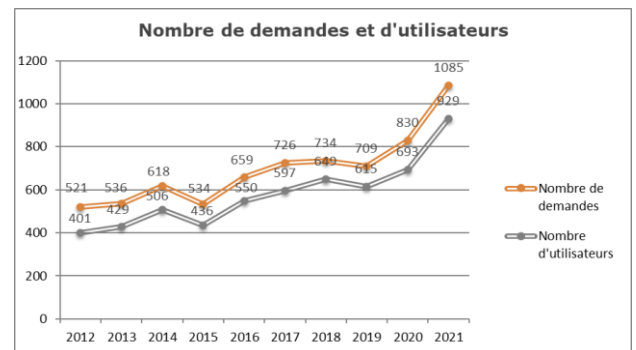
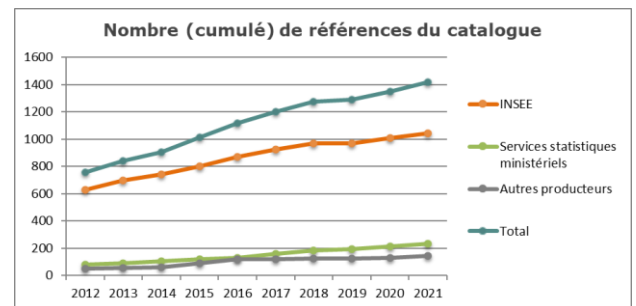
Outre l'activité habituelle de documentation et d'ajouts de nouvelles données, **l'année 2021 a été consacrée à la transformation du portail de diffusion**. Entre autres nouveautés, les jeux de données ont désormais un DOI qui leur est attribué, qui permettra d'améliorer la citation des données. L'ergonomie a été complètement revue, qui permet à travers un site unique d'accéder à l'information aussi bien au niveau des enquêtes, que des séries et surtout d'interroger la base de données au niveau des 400.000 variables qu'elle a documentées. Ce nouveau site (<https://data.progedo.fr>) intègre ainsi les fonctionnalités des trois sites précédents (catalogue, outil de recherche, base de questions).

Archives des Données Issues de la Statistique Publique

L'Adisp a pour mission de diffuser des enquêtes et bases de données produites par la statistique publique, auprès de la communauté scientifique (chercheurs et étudiants) ; et ce afin de développer l'usage des données statistiques dans les sciences sociales. L'Adisp a donc, depuis plus de 25 ans, constitué un fonds de grandes enquêtes et de bases de données intéressant les sciences sociales et continue de l'alimenter grâce à des conventions signées avec l'INSEE, plusieurs services statistiques ministériels et des institutions publiques. Les données diffusées sont accessibles gratuitement à l'ensemble de la communauté scientifique en France comme à l'étranger. L'Adisp apporte également des conseils sur la recherche et l'utilisation des données.

Les fichiers de données diffusés par l'Adisp sont soumis à des règles de diffusion différentes en fonction de leur contenu ou de contraintes imposées par les réglementations en vigueur et les différents producteurs. On distingue ainsi les fichiers dits « standards » et les fichiers dits « production et recherche » (FPR), ces derniers étant produits par la statistique publique à l'intention de la communauté scientifique. En termes de détails, les fichiers « FPR » offrent un niveau intermédiaire d'information entre les données standards anonymisées (plus agrégées) et les données confidentielles plus détaillées, accessibles sous conditions, en particulier auprès du CASD. Pour accéder aux fichiers « FPR », les utilisateurs doivent être habilités par le Comité du Secret Statistique.

EN 2021, l'équipe de l'ADISP a été renforcé par un recrutement d'une ingénieure CNRS. L'activité de l'ADISP est en augmentation constante (nombre de sources mises à disposition, accompagnement des changements législatifs, modernisation de la plateforme de diffusion des données, etc.). Le développement des plateformes universitaires de données a permis de mieux faire connaître ce service. Il en résulte une augmentation forte des utilisateurs (+ 34%), qui provient à la fois d'une augmentation régulière du nombre de chercheurs utilisant les données (+ 10%), mais aussi d'une augmentation très forte des étudiants (+ 60%). Le nombre d'utilisateurs progresse au niveau national et international, mais c'est surtout hors Ile de France que la progression est la plus forte (+ 73%). La distribution disciplinaire n'évolue que peu et reste très inégale, avec presque 50% d'économistes et plus de 25 % de sociologues.



CASD

Né pour répondre aux problématiques d'accès aux données détaillées soulevées par les chercheurs recourant aux analyses quantitatives dans le cadre de leurs travaux, le CASD –conçu par le Groupe des Ecoles Nationales d'Economie et de Statistique (GENES) et l'INSEE– propose un équipement hautement sécurisé permettant de travailler à distance sur des bases de données individuelles détaillées, dans le respect des lois, des exigences de la CNIL et des règles éthiques de confidentialité les plus rigoureuses.

Initialement centré sur les données de la statistique publique, le CASD met aujourd'hui à disposition des données des Ministères de la justice, de l'éducation, de l'agriculture, des finances pour les données fiscales... ainsi que des données entreprise et de Big Data.

Dans le domaine de la santé, certaines données du Programme de médicalisation des systèmes d'information (PMSI) sont mises à disposition via le CASD.

Petit à petit, pour répondre aux besoins de la recherche, le dispositif intègre de nouvelles sources et étoffe le nombre de références accessibles.

On dénombre aujourd'hui 945 projets développés grâce au CASD (<https://www.casd.eu/projets-developpes-sur-le-casd/>)

Autres accès aux données

INSEE

Le partenariat développé avec l'INSEE fait de PROGEDO l'interlocuteur unique pour les chercheurs qui souhaitent accéder à des données qui ne sont pas ouvertes. L'accès peut se faire de deux manières. La première concerne les données confidentielles et repose sur le dispositif du Centre d'Accès Sécurisé aux Données (CASD, voir page 15). La seconde consiste à mobiliser des fichiers dits « de productions et de recherche » (FPR) qui offrent des données pseudonymisées. Les deux procédures requièrent l'aval du comité du secret statistique, l'une gérée directement par le chercheur pour les données confidentielles, et l'autre par l'intermédiaire de l'ADISP (voir page 14) pour les données pseudonymisées (procédure simplifiée).

En outre, PROGEDO propose des produits de l'INSEE pour les chercheurs tels que les « produits sur mesure » (extraction de tableaux à façon) ou le distancier *métric* (outil de mesure géographique de distances entre communes). C'est l'ADISP qui assure ces services.

ICPSR

Basée à l'Université du Michigan, l'Inter-university Consortium for Political and Social Research diffuse plus de 16 000 études issues de différents pays (dont la France) et constitue l'une des plus grandes archives de données en sciences sociales au monde. L'ICPSR offre aussi de nombreuses ressources qui concernent la gestion, l'archivage et la diffusion des données ainsi que leur utilisation. Le site web du consortium propose ainsi de nombreux modules de présentation et de formation. Enfin, l'ICPSR propose aussi chaque été des formations approfondies en présentiel (summer program).

L'accès aux données de l'ICPSR nécessite que le pays fasse partie du consortium. C'est le cas pour la France grâce au financement de PROGEDO pour l'ensemble de la communauté scientifique nationale.

L'accès nécessite l'identification de l'adresse IP de l'ordinateur qui se connectera aux serveurs de l'ICPSR. Cette adresse IP doit être localisée dans un laboratoire de recherche public. Ce sont les représentants officiels auprès de l'ICPSR (OR) qui gère cette identification. En France, les Plateformes Universitaires de Données assurent ce rôle pour PROGEDO.

LIS

Afin de faciliter la recherche comparative au niveau global, le LIS (Luxembourg Income Study Cross National Data Center) harmonise et documente des microdonnées socio-économiques provenant d'enquêtes nationales. Le LIS propose l'accès à deux bases de données distinctes dédiées pour l'une aux revenus et à la consommation et pour l'autre au patrimoine des ménages.

Luxembourg Income Study Database (LIS): la LIS contient des données sur le revenu et la consommation des ménages provenant de plus de quarante pays répartis sur l'ensemble des continents (Europe, Amérique du Nord, Amérique Latine, Afrique, Asie, et Océanie) sur une période de 40 ans, de 1970 à 2010.

Luxembourg Wealth Study Database (LWS): la LWS se focalise sur le patrimoine des ménages. Elle contient des variables sur le système d'imposition, les dettes, les dépenses ou encore la composition des ménages de onze pays occidentaux (Allemagne, Autriche, Canada, Chypre, Etats-Unis, Finlande, Italie, Luxembourg, Norvège, Royaume-Uni et Suède) et du Japon.

L'accès aux données est possible pour les chercheurs français grâce au financement de PROGEDO.

Les données fournies prennent trois formes :

1. des chiffres clefs (indicateurs internationaux basés sur les données harmonisées du LIS). En libre accès.
2. des données agrégées dans un outil interactif en ligne qui permet de comparer et d'extraire des données agrégées, sans aucune connaissance en programmation statistique. En libre accès.
3. des microdonnées accessibles par un accès distant : LISSY est un système en « remote-execution », qui permet de soumettre directement des requêtes statistiques en SAS, SPSS ou STATA. Seuls les résultats finaux sont renvoyés à l'utilisateur. Accès pour les étudiants et les chercheurs enregistrés.

Promotion des données

Pour développer la culture des données et assurer la promotion des données existantes, l'IR* PROGEDO développe une stratégie en 4 axes. Le premier concerne la communication au sens classique et s'appuie sur un ensemble d'outils rédactionnels en ligne. Le second repose sur un maillage territorial au plus près des communautés de chercheurs grâce à un réseau de plateformes universitaires de données. Le troisième s'organise autour de champ thématique, couvert par les départements de l'IR*. Enfin, l'IR* est partenaire du projet structurant LifeObs (financé dans le cadre du PIA).

Communication de l'IR* PROGEDO

Le premier axe de promotion des données repose sur un écosystème lié à internet.

Sites web

Progedo.fr

L'année 2021 a été celle de la refonte complète du site internet de l'IR* (<https://www.progedo.fr/>) afin d'en faire la porte d'entrée principale vers l'ensemble de ses activités. A cette occasion, nous avons sécurisé la navigation des utilisateurs (mise en œuvre du protocole https) et surtout repensé l'intégralité du site.

Le site offre dès sa page d'accueil l'accès à l'ensemble de ses données (Quetelet-Progedo-Diffusion ; CASD ; Enquêtes internationales), présentent ses services (ADISP, PUD, dépôts de données), les éléments de formation et communication mis à disposition (vidéos, tutoriels, ressources et publications) et expose ses départements. Les éléments plus institutionnels (gouvernance, partenaires, informations pratiques et légales) sont accessibles aussi depuis la page d'accueil.

Un système de suivi de la fréquentation a été mis en place pour s'assurer à l'avenir du suivi de son utilisation.

Catalogue de données QPD

2021 a vu la fin de la rénovation du catalogue de données de Quetelet-Progedo-Diffusion, entamé en 2020. Le catalogue présente en un lieu unique les données de tous ses diffuseurs partenaires. Il présente une interface épurée pour trouver rapidement les données. Il remplace ainsi les 3 sites web précédents (« ORDS » pour la recherche de données ; « Base de Question » pour l'interrogation des variables ; « serveurs Nesstar » pour la description des variables).

De nouvelles fonctionnalités ont été implémentées :

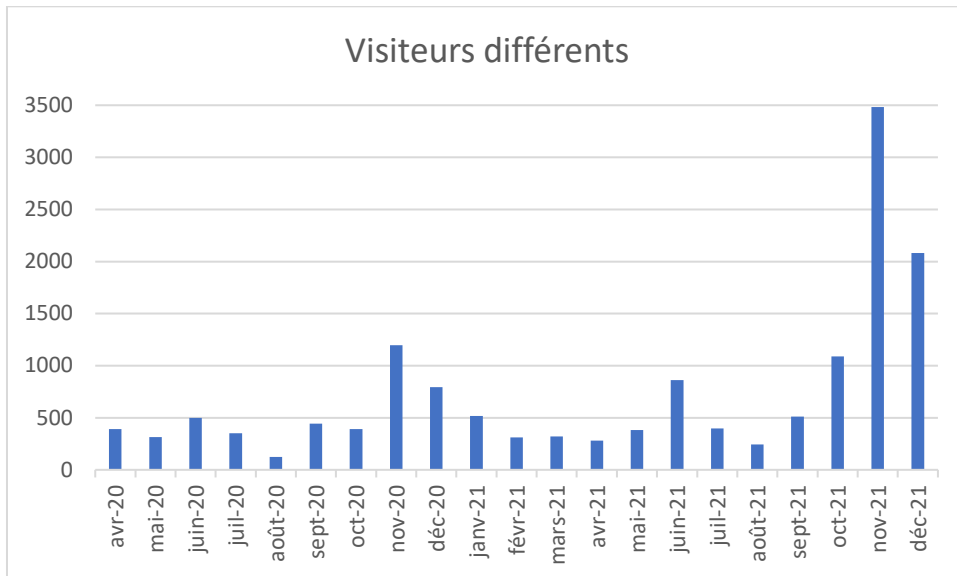
- Un outil de recherche amélioré, qui permet d'explorer les métadonnées des enquêtes et des variables.
- Des filtres pour affiner ses recherches
- Une présentation détaillée des enquêtes et variables
- Un identifiant unique (DOI) pour identifier chaque jeu de donnée

Il offre en outre une compatibilité avec les métacatalogues existants (Isidore de l'IR* 'Humanum; le Cessda Data Catalogue; EOSC).

Carnet hypothèse

L'IR* s'est doté en 2020 d'un « carnet Hypothèses » (mis à disposition par l'IR OpenEdition) sur lequel elle publie son actualité (Évènements, Opportunités, Publications et Ressources). 48 billets ont été publiés en 2021 (contre 39 en 2020). Ce carnet nourrit directement le fil d'actualité du site principal de l'IR*, il permet d'alimenter aussi les sites web des PUD par un système de flux RSS avec filtre personnalisé (par exemple : <https://pud.mmsh.univ-aix.fr/>).

Le carnet a accueilli en moyenne 870 visiteurs différents chaque mois (contre 380 en 2020), avec un pic en novembre-décembre qui correspond à la période entourant les semaines Data-SHS organisées par l'IR*.



canal U

L'IR* dispose aussi d'une chaîne de diffusion sur Canal U (<https://www.canal-u.tv/chaines/progedo>) qui a été entièrement repensée en 2021 (à l'occasion de la refonte de Canal U). Celle-ci permet de valoriser l'ensemble des vidéos produites par l'IR*. Il s'agit de vidéo des conférences et séminaires organisées par l'IR* et ses partenaires. Depuis 2021, une série de tutoriels pour la prise en main d'outils et de méthodes ont été développés par les PUD et son disponible dans un dossier. Par exemple le dossier « travailler avec le logiciel R comprend déjà 6 tutoriels et un cours : <https://www.canal-u.tv/chaines/progedo/travailler-avec-r>)

HAL

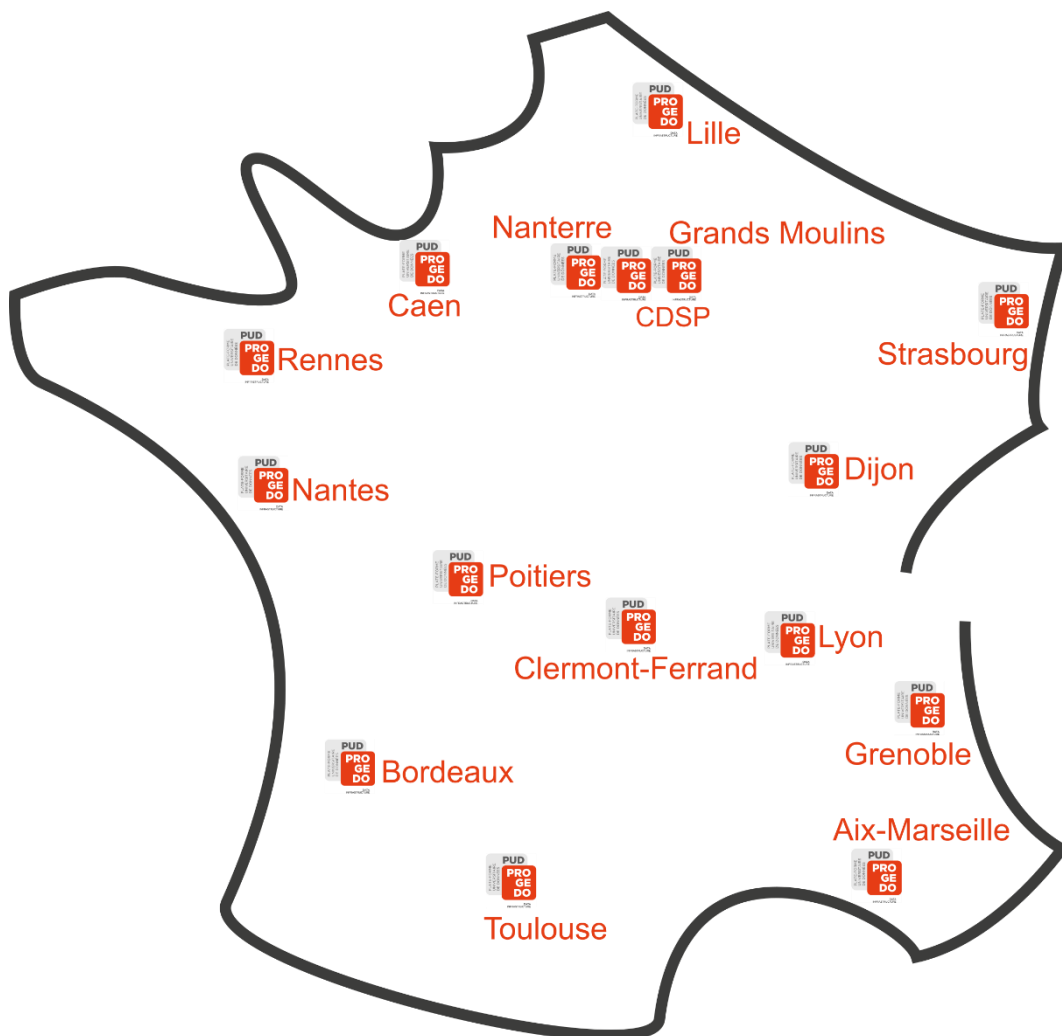
L'IR* dispose depuis 2021 d'une collection dans HAL qui lui permet d'identifier les publications issues de ses productions. La collection permet d'alimenter la rubrique édition du site web principal (<https://www.progedo.fr/formation/edition/>). Son enrichissement se fera au fur et à mesure du temps (l'opération ne peut pas être automatisée), mais elle regroupe déjà plus d'une centaine de références.

Plateformes Universitaires de données

A la suite au rapport Silbermann (1999) sur les sciences sociales et leurs données, la structuration des recherches quantitatives s'appuyant sur la statistique publique et les grandes enquêtes s'est lentement renforcée et a abouti à la mise en place des différents dispositifs en interaction au sein de l'IR* PROGEDO.

3 voies complémentaires avaient alors été tracées, qui organisent l'IR* aujourd'hui. Le regroupement du financement des grandes enquêtes (particulièrement celles inscrites dans des programmes européens de type ERIC) devait permettre leur meilleure valorisation. Ensuite la mise en place d'un dispositif national de diffusion des données permettait d'avoir un guichet unique pour les producteurs et les utilisateurs (dispositif qui prendra le nom de centre Quetelet puis de réseau Quetelet avant de devenir Quetelet-Progedo-Diffusion). Enfin, des dispositifs d'accompagnement locaux étaient présentés comme nécessaire à la diffusion de l'information et l'accompagnement des chercheurs.

C'est dans ce contexte que la première PUD a été construite, à partir d'une initiative locale à Lille en 2001-2002. La seconde PUD a été développée au sein de la MSH-LSE en 2009. PROGEDO a ensuite permis d'amorcer, sur des fonds locaux, les PUD au sein de la MRSH (Caen) et de la MSH Ange Guépin (Nantes) en 2015, puis à la MISHA (Strasbourg) et la MMSH (Aix-Marseille) en 2016. A partir de 2018, le plan national de soutien aux SHS a permis de renforcer l'existant et d'accélérer le développement des PUD. En 2022, il existe désormais 16 PUD couvrant les 15 principaux sites universitaires français.



La coordination par PROGEDO de l'ensemble des PUD permet de conserver en cohérence les objectifs nationaux au niveau local. Elle s'exprime à travers la réunion présentielle des ingénieurs des PUD à deux reprises chaque année, et de réunions plus courtes en visioconférence. Dans un contexte en évolution rapide aussi bien dans le champ scientifique, technique et légal ces réunions sont primordiales. D'autre part, une action concertée est organisée chaque année la seconde semaine de décembre dans l'ensemble des PUD : la « semaine Data SHS ». Durant toute une semaine, les PUD organisent des séminaires et formations sur leur site, mobilisant l'ensemble de la communauté. PROGEDO, au niveau national, assure la cohérence de ces semaines et valorise ces événements. Cela illustre cette manière décentralisée mais extrêmement coordonnée de travailler.

Cette articulation du local et du national a permis cette année la participation des PUD aux ateliers de la donnée partout où les sites universitaires ont souhaité répondre à l'appel à manifestation. Les SHS sont apparus par ce biais comme un champ disciplinaire déjà bien structuré et avancé dans l'accompagnement des chercheurs sur la question de la documentation, la conservation, la diffusion et la réutilisation des données.

L'objectif d'une plateforme est de favoriser le développement des recherches en SHS en s'appuyant sur l'exploitation et le traitement de grands fichiers statistiques :

- Accompagner les travaux des chercheurs expérimentés
- Favoriser la formation des jeunes chercheurs, doctorants et étudiants
- Mettre à disposition de l'enseignement et de la recherche des outils statistiques et des informations adaptées.

En matière de recherche, la PUD apporte un soutien technique individualisé. La PUD n'a pas vocation à effectuer les traitements de données pour les travaux de recherche. Elle a un rôle d'assistance. Elle ne peut se substituer aux chercheurs dans le traitement ou l'exploitation des données.

Accompagnements individuels

En 2022, ce sont 362 personnes (dont 137 étudiants) qui ont été accompagnés individuellement dans leur recherche de donnée, l'aide à leur traitement ou à leur valorisation.

Accompagnements de projets

64 projets ont été accompagnés, le plus souvent dans leur phase d'amorçage, pour préparer la recherche de données ou la mise en place d'un plan de gestion de données. 31 d'entre eux ont été financés par l'ANR, une région ou par d'autres sources.

Formations

Les PUD ont organisé 50 jours de formations ouvertes à tous (méthodes et outils d'analyse de données, recherche de données) et animé 25 formations doctorales. Elles sont intervenues à 34 reprises dans des diplômes de masters ou licence.

Séminaires et conférences

13 séminaires, journées d'études ont été organisés (journée d'étude sur l'enquête ISSP, séminaire sur la nomenclature des PCS, journée sur les données en santé, etc.)

Semaine Data SHS

Chaque année depuis 2019, les PUD organisent simultanément une semaine de promotion des données quantitatives en SHS lors d'une semaine nommée « Semaine Data SHS ». C'est l'occasion d'organiser des ateliers méthodologiques, des présentations de recherche et des formations autour

des données (grandes, enquêtes, statistiques publiques) et de leur usage (plan de gestion des données, RGPD, etc.).

La semaine Data SHS a touché plus de 700 personnes en 2021.

Ateliers de la donnée

Partout où les sites universitaires ont souhaité répondre à l'appel à manifestation d'intérêt « ateliers de la donnée » lancés par le Mesri, les PUD ont participé. Au total, 10 des 16 PUD sont engagés dans le premier AMI Atelier de la donnée.

Les départements

Comme mentionné précédemment, l'IR* PROGEDO s'organise thématiquement autour de grands départements chargés d'animer des communautés de recherche transverses aux disciplines. Ces départements œuvrent en synergie et interagissent avec les partenaires de l'IR* au sein des SHS (équipes des enquêtes internationales, équipe de la diffusion nationale (QPD), équipes locales des PUD) et au-delà. Chaque département a développé des objectifs spécifiques correspondant aux enjeux relevés pour leur thématique, en agissant pour valoriser les données existantes et travaillant à préserver ou créer les données nécessaires.

Méthodes et statistiques publiques

La mise à disposition de données est un enjeu essentiel de la science ouverte aujourd'hui. Mais les données doivent être « aussi ouvertes que possible et aussi fermées que nécessaires » pour protéger les personnes. Entre loi pour une république numérique et RGPD, le département « méthodes et statistiques publiques » s'emploie à proposer aux chercheurs des données non ouvertes au public, en respect de la loi et sous condition de confidentialité. Il met à disposition des chercheurs des fichiers pseudonymes et promeut l'utilisation des données individuelles grâce au Centre d'Accès Sécurisé aux données.

Le travail est le fruit de partenariats avec les producteurs de la statistique publique (en premier lieu l'INSEE et les services statistiques ministériels).

De la même manière, la diffusion des données de la recherche est aussi une préoccupation du département. La diffusion de données ne peut cependant pas se faire sans qu'elles soient accompagnées d'une documentation qui permettra leur réutilisation. Le département travaille donc à promouvoir les bonnes pratiques et accompagne pour cela l'ensemble des communautés de recherche à travers la promotion de standards (DDI) et la formation.

Le département est animé par Muriel Roger (professeur à l'université Paris 1 Panthéon Sorbonne, membre du laboratoire CES).

Opinions et enquêtes et sociopolitiques

PROGEDO soutient la participation de la France dans la production de grandes enquêtes pluridisciplinaires d'intérêt national. L'IR* s'implique à ce titre dans les consortia constitués comme ERIC autour des enquêtes européennes. Elle soutient aussi les enquêtes internationales dont le périmètre géographique dépasse le continent européen – et qui n'ont donc pas le statut d'ERIC –, mais qui sont néanmoins stratégiques pour les communautés scientifiques en SHS.

Le département Opinions et Enquêtes Sociopolitiques, animé par Frédéric Gonthier (professeur à Université Grenoble Alpes, Sciences Po Grenoble, Pacte-CNRS), vise à consolider la présence de PROGEDO dans le paysage national en tant que centre de compétences sur les grandes enquêtes internationales, selon un modèle analogue à celui des groupes de recherche International Survey Programs de GESIS en Allemagne ou de Valeurs et attitudes de FORS en Suisse.

Les objectifs du département s'organisent autour des trois domaines d'activités suivants :

- Impulser et accompagner la dynamique de structuration des grandes enquêtes sociopolitiques internationales (ESS, EVS et ISSP), en contribuant à leur pérennisation et en facilitant les synergies (entre enquêtes et avec les dispositifs nationaux de collecte de données).
- Aider à coordonner et à consolider les activités de valorisation des grandes enquêtes sociopolitiques internationales, afin d'améliorer leur visibilité académique et extra-académique.

- Promouvoir les méthodes, notamment innovantes, liées au recueil et à l'analyse quantitative des grandes enquêtes sociales internationales, en intensifiant les échanges scientifiques entre les programmes d'enquêtes et en les rapprochant des actions de formation déjà existantes.

Santé

Les données de santé sont d'autant plus diverses que l'on appréhende la santé en lien avec le bien-être et la qualité de vie des personnes comme peuvent le faire les SHS. Au sens strict, elle font l'objet d'une cadrage et d'une réglementation précise en terme de diffusion et d'accès (Il existe d'ailleurs une structure dédiée à leur diffusion: le système national des données de santé). Au sens élargi, elle peuvent concerner l'étude des modes de vie et le ressenti qu'en ont les personnes. L'ambition de ce département est à la fois de favoriser l'accès aux données de santé (au sens strict) de façon élargie au sein des communautés de chercheurs en SHS et de promouvoir la production et le partage de données de santé au sens élargi du terme.

Le département est animé par Philippe Terral (professeur à l'Université Paul Sabatier de Toulouse, et membre du laboratoire CRESCO).

Données historiques économiques et financières

Les bases de données historiques demandent des efforts et du temps pour être construites et constituent un patrimoine qu'il faut préserver et mettre en valeur. L'ambition de ce département est de faire vivre ces bases de données dont certaines sont inédites.

Leur préservation et leur large diffusion sont rendues possibles grâce au soutien de l'équipe de l'Adisp-Progedo. Pour cela, les données doivent être documentées et Progedo peut aussi accompagner les chercheurs dans ces pratiques.

Le département est animé par Sandra Brée (chargée de recherche CNRS au LARHRA).

Population et parcours de vie

Le département « Population et parcours de vie » vise d'une part à soutenir les grandes enquêtes françaises et internationales qui s'intéressent à la famille au sens large (individus, ménages, fécondité, etc.), aux parcours de vie et plus généralement aux questions de population. D'autre part, il cherche à promouvoir l'usage de ces données par le plus grand nombre. Dans ce cadre, le département souhaite mieux faire connaître ces données et développer leur usage dans le champ pluridisciplinaire des *population studies* et au-delà.

Pour ce faire, le département travaille en étroite collaboration avec le projet LifeObs porté par l'INED, soutient ou promeut les enquêtes « Familles », « GGP » (aussi appelée ERFI pour son volet français), « Famille et employeurs », « FECOND », « AMP sans frontières », « Eurocohort », « Envie » et « SHARE ».

Le département est animé par Pascal Sébille (Maître de conférences à l'université de Rennes 2, UMR ESO CNRS) qui s'appuie sur le réseau des Plateformes Universitaires de Données et plus particulièrement sur la PUD de Strasbourg et celle de Bordeaux.

Ce département a été conçu en 2021 et démarrera en septembre 2022.

Life Obs

L'ambition de l'Observatoire des parcours de vie est de fournir aux sciences sociales les infrastructures nécessaires pour collecter et diffuser des données de qualité sur des sujets fondamentaux concernant la dynamique des parcours de vie, l'articulation famille-travail et les tendances démographiques. La France dispose d'un riche patrimoine d'enquêtes quantitatives à grande échelle mais le contexte actuel de changements sociaux, économiques et politiques rapides appelle un renouvellement des investissements. Pour que la France conserve un rôle de premier plan dans les programmes d'enquêtes internationaux, LifeObs se donne une double mission.

1. LifeObs documentera et diffusera les recherches existantes en améliorant la visibilité des enquêtes françaises. L'objectif est de valoriser les investissements précédents en s'assurant que les données françaises sont largement accessibles et utilisées par les chercheurs nationaux et internationaux ainsi que par les décideurs politiques. Cet objectif se décline en trois actions distinctes. Premièrement, LifeObs traduira les enquêtes existantes et leur documentation afin de les faire connaître et de les rendre exploitables par les chercheurs étrangers. Deuxièmement, il renforcera la diffusion des enquêtes françaises à travers des plateformes de données internationales et au sein de réseaux internationaux. Troisièmement, il formera à l'analyse des enquêtes quantitatives et contribuera ainsi à une montée en compétences au sein des sciences sociales françaises.

2. LifeObs investira et innovera dans la collecte de nouvelles données. Il développera l'infrastructure nécessaire pour doter les chercheurs et les décideurs politiques d'enquêtes de grande qualité et à grande échelle sur un large éventail de sujets liés aux comportements familiaux et à la dynamique des populations, couvrant toutes les étapes du cycle de vie. Ces enquêtes, qui forment un ensemble cohérent et dont beaucoup sont longitudinales, porteront sur les principales questions scientifiques et politiques émergentes, telles que la baisse de la fécondité, la procréation médicalement assistée, la conciliation famille-travail, les comportements à risque chez les jeunes, le vieillissement de la population. Cet investissement est essentiel pour permettre à la France de participer à des programmes d'enquêtes internationales tels que le prochain cycle du GGP, SHARE ou la première cohorte européenne d'enfants GUIDE-EuroCohort.

Cet observatoire multisite repose sur une coopération nationale entre des institutions clés dans le domaine des études sur les parcours de vie, la famille et la population. L'objectif est de mutualiser à la fois les ressources (données et documentation) et les compétences (accès et analyse des enquêtes) afin de les rendre largement accessibles à la communauté scientifique. L'Ined est le principal centre de recherche sur la population en France et l'un des centres internationaux de référence en démographie. L'INSEE a une longue tradition de réalisation d'enquêtes à grande échelle auprès des ménages sur des thèmes liés à la famille. La très grande infrastructure de recherche PROGEDO est en charge de la participation française aux grandes enquêtes internationales, de la diffusion des données et de la formation des utilisateurs. LifeObs associera ces institutions à trois grandes universités françaises. L'Université Paris Dauphine - PSL réalise l'édition française de l'enquête SHARE. Les universités de Bordeaux et Strasbourg accueillent les principaux départements d'études démographiques en France et fournissent un soutien crucial aux utilisateurs de données notamment via des Plateformes universitaires de données. En fédérant ces acteurs, LifeObs renouvellera et renforcera radicalement la production, la diffusion et la formation à l'utilisation des données françaises. Cet effort aura un impact immédiat sur la production scientifique et servira de référence pour les politiques publiques, les médias et la société civile.

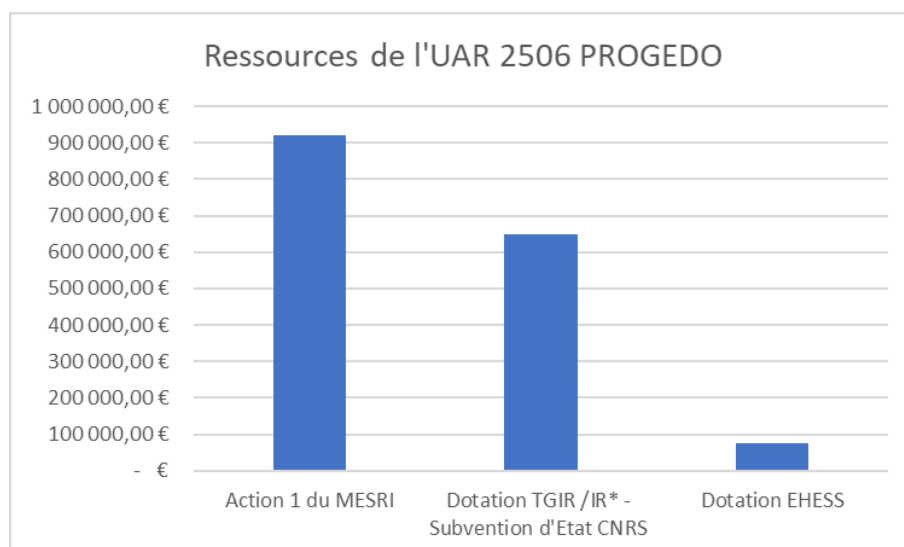
Le projet Life Obs a démarré en novembre 2021 et bénéficie d'un financement du programme d'investissements d'avenir

Bilan financier

Ressources

Les ressources financières en 2021 s'élevaient à 1.647.000 €, dont :

- 56% de subventions du MESRI.
- 44 % de dotations d'état (76.400 € de l'EHESS et 649.100 € du CNRS)



Les ressources de l'IR* PROGEDO sont composés de :

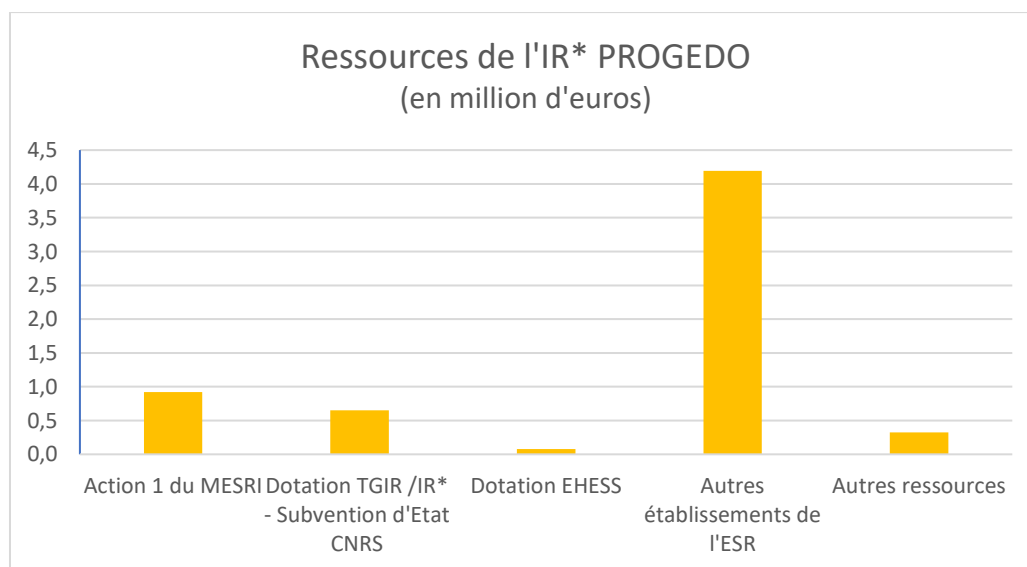
L'action 1 du MESRI à hauteur de 920.000€ (15% du total des ressources)

La dotation du CNRS à l'UAR à hauteur de 650.000€ (11% du total des ressources)

La dotation de l'EHESS à l'UAR à hauteur de 76.000€ (1% du total des ressources)

L'engagement des autres établissements de l'ESR participant à l'IR* à hauteur de 4.195.000€ (68% du total des ressources)

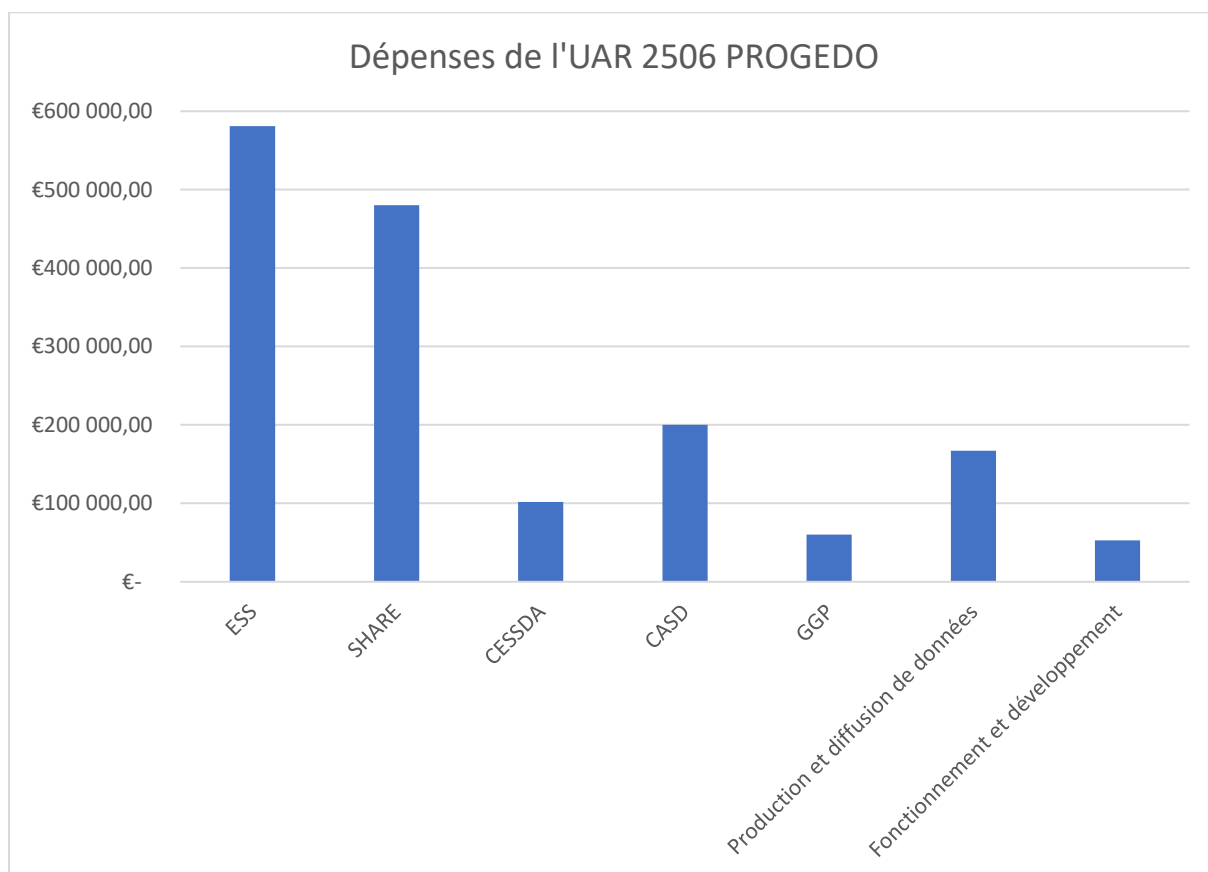
Les autres ressources (partenaires institutionnels hors ESR) à hauteur de 325.000€ (5% des ressources)



Dépenses

Les dépenses de 2021 se répartissent en trois postes distincts :

- Le soutien aux grandes enquêtes européennes ESS, SHARE et GGP (68,2% des dépenses) dont le montant s'élève à 1.121.000€
- L'accès aux statistiques publiques et à des bases de données internationales (31,8,2%) pour 522.000€, qui comprennent :
 - o la sécurisation de cet accès (200.000 € versés au CASD, soit 12,2% des dépenses)
 - o les données fournies par l'INSEE (diffusées par l'ADISP), le LIS Cross National Data Center et l'ICPSR (Inter-University Consortium for Political and Social Research) et le soutien aux productions nationales (167.000€, soit 10,2% des dépenses)
 - o la contribution à l'ERIC CESSDA (102000 €, soit 6,2% des dépenses)
- Les frais de fonctionnement de l'UAR 2006 Progedo (3,2%), soit 53000 €, qui couvrent les dépenses de maintenance des sites de Progedo, les déplacements des membres de l'équipe (participation à des manifestations scientifiques nationales et internationales, coordination des PUD) et le soutien ponctuel à des formations



Les dépenses de l'IR* se partagent en dépenses de fonctionnement (49,4% du total) et dépenses de personnel permanent (35,3% du total) et non permanent (15,3% du total).

Dépenses de l'IR* PROGEDO (en million d'euros)

