

Rapport d'activité PUDC 2013

Avril 2013

Rapport rédigé par Clotilde Lemarchant, MCF de sociologie à l'UCBN et responsable de la PUDC et Cyril Coinaud, ingénieur d'étude statisticien à la PUDC.

Sommaire

Objet et fonctionnement de la PUDC.....	1
Un service de soutien à la recherche quantitative aujourd'hui bien identifié.....	3
Bilan au sujet des utilisateurs et utilisatrices de la PUDC	4
Un travail en réseau	9
Adossement à l'enseignement supérieur.....	10
Adossement à la recherche	11
Adossement à des missions internes à l'université.....	11
Projets, perspectives et questions	13
Bilan Financier	14
Annexes	15

Objet et fonctionnement de la PUDC

Objet

Une plateforme universitaire de données vise à faciliter l'accès aux grandes enquêtes et données statistiques intéressant les sciences humaines et leur utilisation.

Outil de proximité à l'égard des étudiants, enseignants et chercheurs de l'université de Caen Basse-Normandie dans toutes les disciplines des sciences humaines, la PUDC offre ses services dans la recherche de fichiers adaptés aux recherches, les démarches administratives avec le réseau Quetelet et la récupération et l'installation des bases de données. Elle propose un accompagnement méthodologique et des formations adaptées, de la conception des questionnaires jusqu'au traitement statistique. La PUDC permet de connaître les organismes, les grandes enquêtes les plus appropriées à une recherche, les démarches pour obtenir les fichiers, la déontologie à respecter dans l'utilisation de ces fichiers. Elle propose un accompagnement personnalisé en traitement de données

et analyse statistique : interprétation, description d'une distribution, tests statistiques, méthodes d'analyse, mais aussi gestion de fichiers, navigation sur le WEB, utilisation des logiciels statistiques, transfert des données d'un logiciel à un autre.

Elle a enfin pour fonction d'archiver de façon pérenne les données des recherches réalisées dans le cadre de l'université de Caen, quantitatives ou qualitatives.

Contexte national

La PUDC est partenaire du GIS Réseau Quetelet, Groupement d'intérêt scientifique, dispositif national d'archivage, de valorisation et de diffusion des données pour la recherche qui met à disposition gratuitement des données à destination de la recherche.

Les plateformes universitaires de diffusion des données en sciences humaines et sociales font partie du dispositif national de mise à disposition de la recherche des grands fichiers d'enquête produits par des organismes publics ou parapublics (INSEE, INED, Ministères, CNRS, Centres de recherches universitaires ou publics, etc.).

Avec le Comité de concertation pour les données en sciences humaines et sociales et le Réseau Quetelet, les plateformes de données forment le troisième élément d'une infrastructure nationale visant à développer l'utilisation des enquêtes et des grands fichiers de données en sciences humaines et sociales. La première plateforme de ce type a été créée à Lille en 2002, la seconde à Lyon en 2008, la troisième à Caen en 2011.

Financement et fonctionnement : un soutien complet de l'université de Caen UCBN

D'abord financée dans le cadre d'un PPF du précédent quadriennal qui a pris fin en décembre 2011, la PUDC est désormais entièrement financée et soutenue par l'Université de Caen UCBN depuis le 1^{er} janvier 2012.

La PUDC a débuté en janvier 2011, avec un ingénieur d'étude, Cyril Coinaud, financé pour un an sur la fin d'un financement PPF, et une responsable, Clotilde Lemarchant, MCF de sociologie à l'UCBN.

Elle est localisée à la Maison de la recherche en sciences humaines (MRSH) de l'université de Caen Basse-Normandie, au bureau n° 109 (1^{er} étage). Un site internet est disponible à l'adresse suivante : <http://www.unicaen.fr/recherche/mrsh/pudc>

- Cyril Coinaud est ingénieur d'études, Tél. : 02 31 56 61 43 ; Courriel : pudc@unicaen.fr et cyril.coinaud@unicaen.fr.

- Clotilde Lemarchant, est la responsable de la plate-forme ; Courriel : clotilde.lemarchant@unicaen.fr

La MRSH Normandie Caen, lieu et moyen de mutualisation de la recherche en sciences humaines et sociales, notamment au travers de l'ingénierie, des centres de ressources et d'équipements, a soutenu la PUDC dès sa création en lui procurant des locaux, finançant ses premiers frais, et lui

laissant libre accès à ses services mutualisés, notamment de communication. Par définition, la PUDC est au service de la recherche pluridisciplinaire et participe modestement à la mutualisation de ces services.

Un comité de pilotage a été constitué, composé d'une quinzaine de membres :

- La secrétaire générale du Comité de Concertation pour les Données en SHS
- Un-e représentant-e des membres du réseau Quetelet
- Un-e représentant-e de la présidence de l'université de Caen
- Un-e représentant-e du CNRS en Région
- Le représentant de la MRSH de Caen
- Les représentant-e-s des laboratoires de recherches en SHS de Basse-Normandie
- Les représentant-e-s des responsables de masters
- Les représentant-e-s des écoles doctorales
- La responsable et l'ingénieur de la plateforme

Un service de soutien à la recherche quantitative aujourd'hui bien identifié

Après deux ans de fonctionnement, la PUDC est aujourd'hui bien identifiée par bon nombre de chercheur-e-s et d'étudiant-e-s.

Un important travail de communication

Informer ...

Il faut continuer de réaliser un travail de communication constant, permanent auprès de tous les collègues de l'université de Caen

- Présentation à l'Ecole Doctorale Homme, Sociétés, Risques, Territoire (ED 556) et Ecole Doctorale Histoire, Mémoire, Patrimoine, Langage (ED 558) le 15/11/2012 ainsi qu'à l'Ecole Doctorale Economie-Gestion Normandie (ED 242) le 30/11/2012.
- Présentation de la PUDC lors de réunions de laboratoires et d'UFR
 - o Présentation au master de Géographie le 26/09/2012.
 - o Présentation à l'AG d'ESO le 15/10/2012
 - o Présentation au Laboratoire CESAMS le 28/03/2013
- Communication informelle (rappels réguliers dans le cadre de la vie sociale informelle de l'université).
-

S'informer ...

Journée INED, Récolter, explorer les données avec R le 3/12/2012.

Dans le cadre du séminaire «Les rencontres de la statistique appliquée », l'Institut National d'Études Démographiques organisait un colloque d'une demi-journée sur l'utilisation du logiciel R dans divers champs de la statistique : analyses multivariées, analyses de données géographiques et de données textuelles par la présentation de quelques packages. Le logiciel R tend à supplanter les autres logiciels d'analyse statistique dans la recherche publique mais aussi au sein de la recherche privée. En effet, le développement rapide de packages, l'augmentation quantitative et qualitative des documentations associées à ces packages, l'argument économique (R est un logiciel libre) et les efforts faits dans l'ergonomie de la visualisation des résultats en sont ses principaux points forts.

Journée Produit sur Mesure de l'état-civil 17/12/2012 à l'INSEE Basse-Normandie.

L'Insee Basse-Normandie organisait dans ses locaux à Caen, une présentation pour le Réseau Quetelet pour les données de l'Etat-Civil et les différentes possibilités. Cette présentation était, outre l'occasion de découvrir l'horizon qu'offrent ces données, un moment pour rencontrer les membres de l'équipe de l'ADISP qui s'étaient déplacés en nombre pour cet événement ainsi que les personnels de l'INSEE local ce qui peut être important pour de futures collaborations.

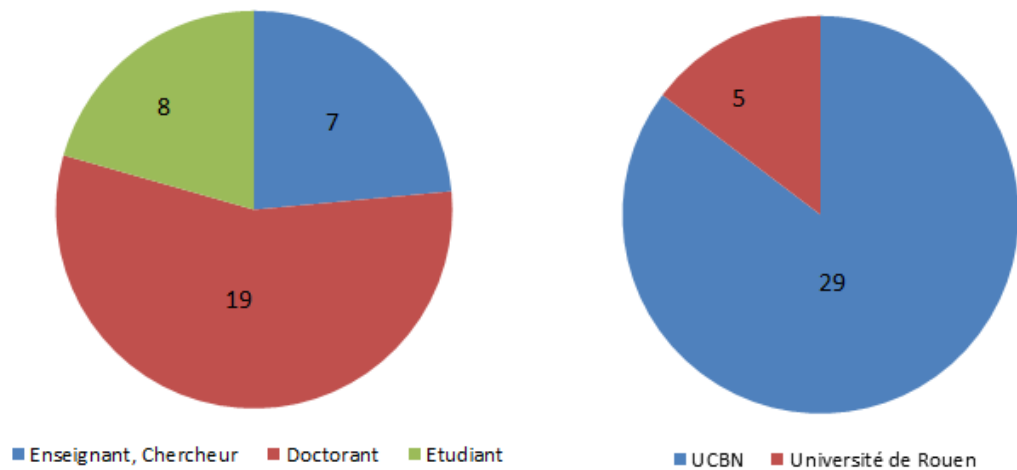
Bilan au sujet des utilisateurs et utilisatrices de la PUDC

1) Le bilan de l'année 2012.

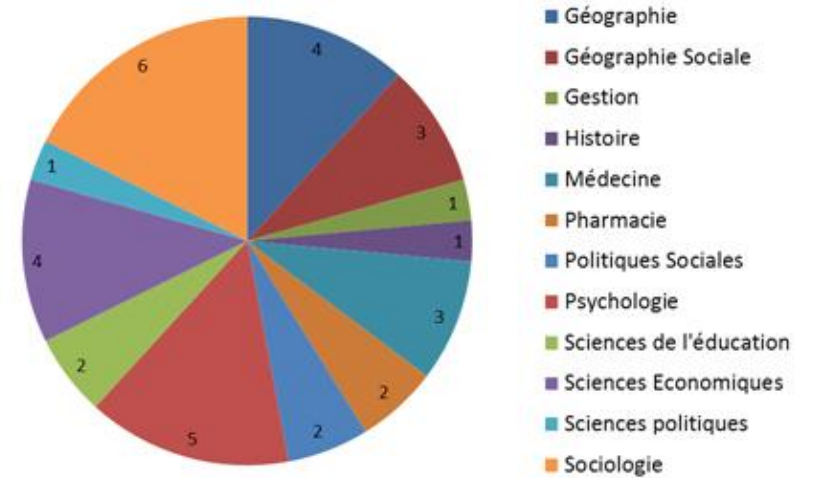
Les nouveaux utilisateurs de la plateforme entre les 2 rapports d'activités (c'est-à-dire entre le 13 mars 2012 et le 1^{er} avril 2013) sont au nombre de 34. Les publics sont assez variés avec des contacts plus fréquents chez les doctorants. Les disciplines sont aussi très diversifiées mais contrairement au dernier rapport d'activité, la couverture au sein des disciplines s'est densifiée : si la sociologie, la géographie sociale et les sciences économiques sont encore présentes dans l'activité de la plateforme, d'autres disciplines ont eu recours à ses services : la Géographie physique, la psychologie, ...

Autre nouveauté cette année : même si les utilisateurs de la PUDC sont essentiellement issus de l'université de Caen, une part non-anecdotique de rouennais (5), doctorants pour la plupart, nous ont sollicités.

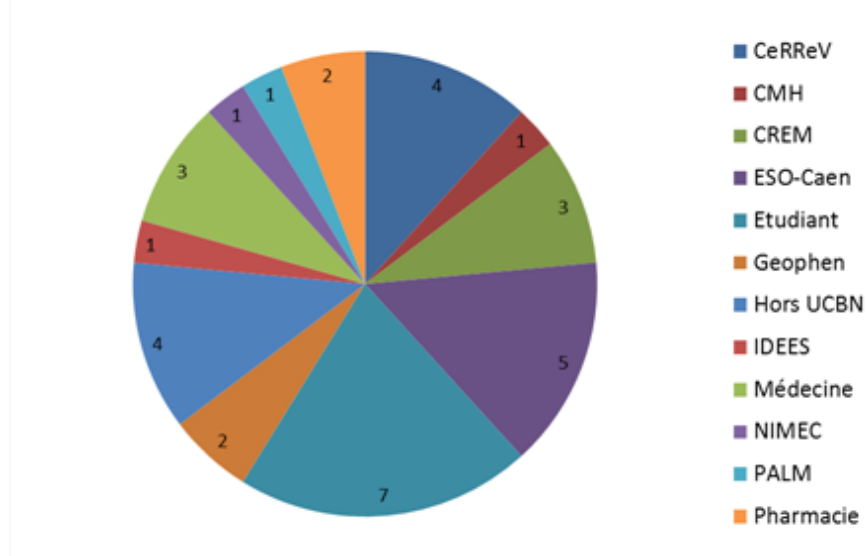
La diversification des disciplines a pour corollaire une meilleure répartition, cette année, des laboratoires d'accueil des utilisateurs de la plateforme même si ESO-Caen, le CERReV et le CREM sont encore cette année, les plus gros pourvoyeurs d'utilisateurs. Il est à noter que les étudiant-e-s en master n'ont pas de laboratoire de rattachement.



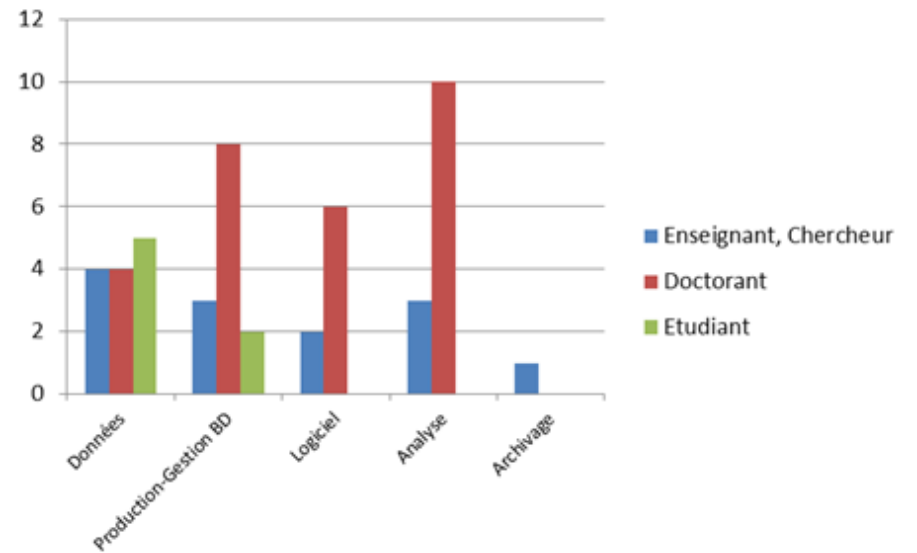
Graphique 1 : répartition des utilisateurs de la PUDC par statut pour l'année 2012.



Graphique 2 : répartition des utilisateurs de la PUDC par discipline pour 2012.



Graphique 3 : répartition des utilisateurs de la PUDC par laboratoire pour 2012.



Graphique 4 : répartition des besoins des utilisateurs de la PUDC pour 2012.

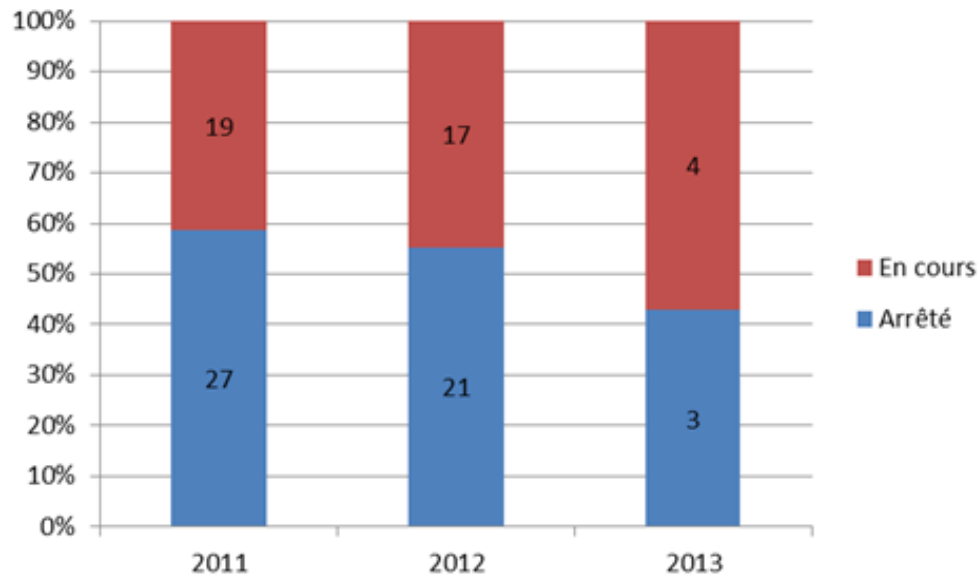
2) Depuis le début de la création de la plateforme.

Dans la partie précédente, nous avons pu faire un état des lieux des flux d'entrants pour l'année d'activité (de mi-mars 2012 jusqu'à début avril 2013). Maintenant, nous nous intéressons un peu plus au « stock » des utilisateurs. C'est-à-dire ceux qui ont sollicité les services de la PUDC pour la première fois durant la première année et qui continuent de la fréquenter. Ensuite nous regarderons aussi les évolutions en termes de public années après années.

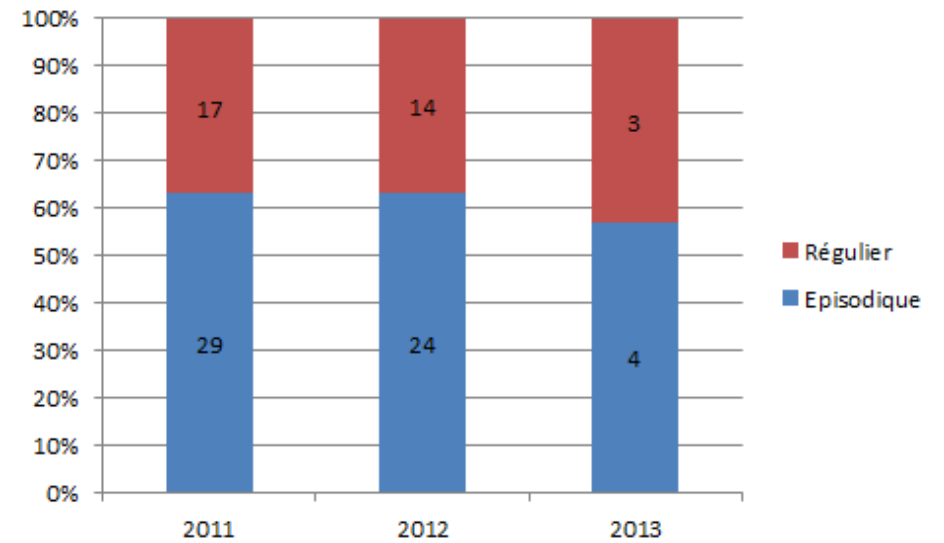
Sur les 46 personnes qui ont demandé de l'aide pour la première fois auprès de la PUDC en 2011, 19 continuent à avoir recours à ses services. Pour les 38 nouveaux utilisateurs durant l'année 2012, 17 entretiennent toujours des liens avec la plateforme. En 2013, déjà 3 demandes concernaient des demandes d'informations autour de logiciels d'analyse de données textuelles auxquelles les réponses ont été apportées.

On peut faire une typologie des utilisateurs en distinguant les « réguliers », qui utilisent à plusieurs reprises les services, et les « épisodiques » qui ne viennent (du moins à ce jour) que pour une demande ponctuelle. On dénombrait 17 « réguliers » en 2011, 14 en 2012 et 3 en (mars) 2013, sachant que la grande majorité des usagers qui ont des entrevues régulières sont des « abonnés » sur la longue durée et viennent depuis le début.

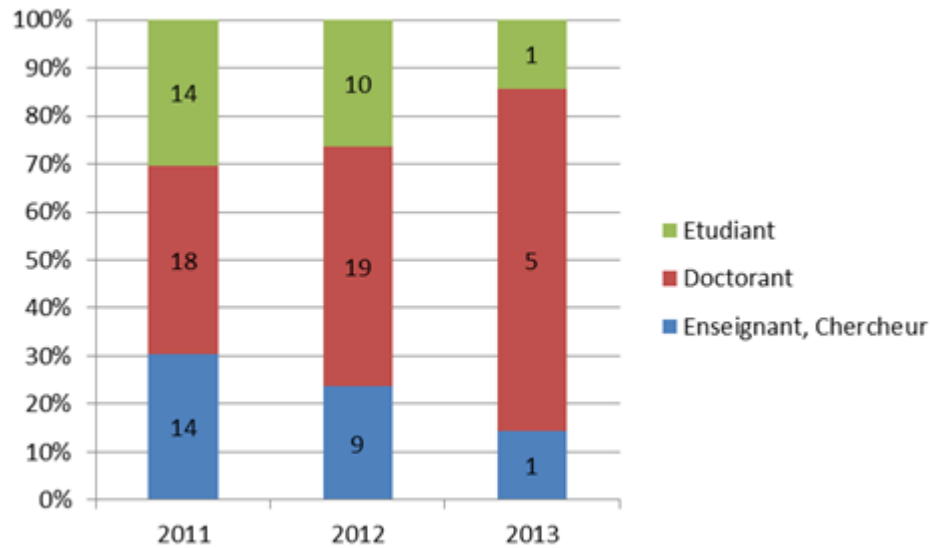
Si les doctorants tendent à avoir une part de plus en plus importante au sein des utilisateurs de la PUDC, ils sont la moitié à entretenir encore ce contact. Les enseignants-chercheurs et les chercheurs ainsi que les étudiants en master sont légèrement moins représentés au fil du temps mais le premier groupe est un groupe plus « d'abonnés » que ne peut l'être celui des étudiants.



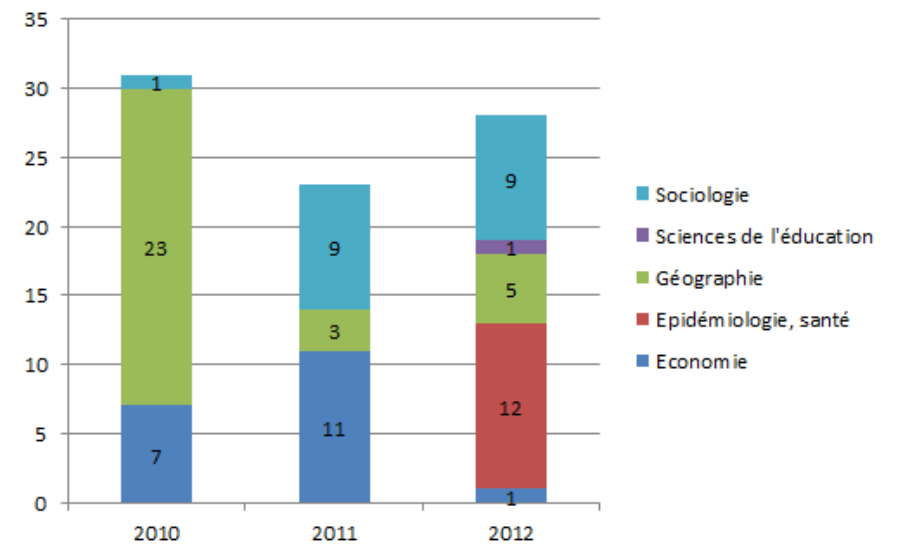
Graphique 5 : répartition des utilisateurs de la PUDC par arrêté selon l'état de leur dossier



Graphique 6 : répartition des utilisateurs de la PUDC par années



Graphique 7 : répartition des utilisateurs de la PUDC par années selon leur statut



Graphique 8 : répartition des commandes auprès du réseau Quetelet par années

Les commandes de fichiers auprès du Réseau Quetelet.

Les commandes du nombre de fichiers de données auprès du Réseau Quetelet (cf graphique 8) sont relativement stables en nombre depuis 2010 : (31 en 2010, 23 en 2011 et 28 en 2012). Si l'on s'intéresse au nombre de fichiers par discipline, on peut remarquer une baisse de demandes en géographie (mais en 2010, il y a eu une grosse commande de la part d'un professeur en géographie physique sur une recherche particulière) et une forte demande le domaine de la santé en 2012 mais là encore, il s'agit de l'effet d'une demande particulière concernant plusieurs fichiers. On peut cependant, noter une augmentation de demandes en sociologie mais une diminution en économie.

Enquêtes demandées	2010	2011	2012	Total
Base centrale de Scolarité	2	7		9
Cadre de vie et sécurité (CVS) - 2010			1	1
Conditions de vie des étudiants			2	2
Correspondances Adresses / Ilots; IRIS / Ilots			2	2
Démographie des entreprises et des établissements - 2006	1			1
Enquête Emploi (en continu)	12			12
Enquête emploi du temps			2	2
Panel de la DEPP 1995 (panel + EVA)		2	2	4
Enquêtes Permanentes sur les Conditions de Vie des ménages			2	2
Formation, qualification professionnelle (FQP)	1			1
Enquêtes Génération	1	5	1	7
Impôt sur le revenu dans les communes de France		3		3
Mode de garde et d'accueil des jeunes enfants - 2007		1		1
Panel des médecins généralistes (PMG) - 2007-2009			1	1
Recensement de la population	13	3	10	26
Revenus fiscaux localisés des ménages			2	2
Structure des salaires (ESS) - 2002	1			1
Système d'information sur les nouvelles entreprises (SINE)		2		2
TIC auprès des ménages (TIC) - 2010			1	1
Transports et déplacements (ENTD) - 2008			1	1
Univers culturels des enfants et adolescents			1	1
Total	31	23	28	82

Tableau de référencement des fichiers demandés auprès du Réseau Quetelet par les utilisateurs

Les enquêtes les plus demandées sont celles qui concernent le recensement de la population (26 fichiers) ; viennent ensuite celles de l'enquête emploi en continu (12) et enfin celles qui portent sur la scolarité (Base centrale de Scolarité : 9, Génération : 7 et Panel DEPP : 4).

On peut remarquer aussi que si le nombre de fichiers demandés est relativement stable sur les 3 années, durant l'année 2012, une diversification de la nature des demandes semble avoir lieu.

Un travail en réseau

Le Réseau Quetelet CCDSHS

Le 1^{er} mars 2013 s'est tenue à l'EHESS, avenue de France, une réunion des équipes participant aux opérations inscrites dans le périmètre du GIS réseau Quetelet. Roxane Silberman y a présenté le nouveau cadre institutionnel résultant de la création de l'UMS PROGEDO-Réseau Quetelet dans le cadre de l'inscription du TGIR PROGEDO au CNRS, le périmètre de cet ensemble avec les deux dimensions, grandes enquêtes pour la recherche et archivage/documentation/ diffusion, la nouvelle gouvernance de cet ensemble avec le CCDSHS, le GIS réseau Quetelet et l'UMS, et ses ressources. Le rôle des plates-formes y est réaffirmé : il s'agit de relayer l'action du réseau Quetelet au plus proche des utilisateurs et leur apporter un soutien pour rechercher et utiliser des données, l'objectif étant de faire croître le nombre d'utilisateurs de données disponibles.

Liens avec les autres PUD

- La PUDC travaille en partenariat avec les deux autres plateformes. Notamment avec la PUDL de Lille, particulièrement dynamique et la plus ancienne puisqu'elle a été créée en 2003. Des contacts sont régulièrement pris pour des conseils ou des échanges techniques.
- Les Centres d'Accès Sécurisés Distants aux Données (CASD), ou comment accéder de manière sécurisée aux données confidentielles de la statistique publique »

La Plateforme Universitaire des Données de Lille proposait aux chercheurs et doctorants en SHS ainsi qu'aux étudiants de master un séminaire d'information sur les CASD le 29/11/2012. Cyril Coinaud a assisté à cette séance. Créés en 2009, ces systèmes permettent de travailler à distance sur des données confidentielles de la statistique publique en utilisant un matériel sécurisé. Données très détaillées, appariements de fichiers, les possibilités sont nombreuses mais encore méconnues. Une première intervention de Michel Isnard, du service juridique de l'Insee, exposait les questions juridiques de la diffusion de données très détaillées ainsi que la politique de son établissement en ce qui concerne la diffusion des données en fonction des publics concernés et du détail des données demandées : une mise à disposition de certaines données au grand public directement sur le site de l'Insee, des fichiers anonymisés de certaines enquêtes accessibles via le réseau Quetelet, des extractions de fichiers sur enquêtes (Produits sur Mesure, gratuites pour les utilisateurs du réseau Quetelet) et enfin les données les plus détaillées (voire individuelles) via le Centre d'Accès aux Données Sécurisées (CASD). Dans un second temps, Kamel Gadouche, porteur du projet CASD et membre du GENES, a expliqué en détails la procédure de demande d'accès aux données du CASD et s'est appuyé sur une démonstration de la SD-box pour en montrer le fonctionnement et les possibilités. Dans une troisième intervention, Olivier Godechot de l'ENSAE, témoignait en quoi l'accès au CASD avait eu pour lui un impact important pour l'accès certaines données financières (études des hauts salaires). Enfin, Raphaëlle Fleureux, membre de l'équipe ADISP

du Centre Maurice Halbwachs, a montré les différentes enquêtes disponibles et les améliorations du site de l'ADISP.

Le RN MSH (réseau national des maisons des sciences humaines)

- Le 28 septembre 2012, s'est tenue à Paris, à l'EHESS avenue de France, une réunion des plates-formes technologiques hébergées dans les MSH. Le RNMSH, réseau national des maisons des sciences humaines était à l'initiative de cette réunion. Clotilde Lemarchant y a présenté la PUDC.
- Les 6 et 7 décembre 2012 s'est tenu le congrès national des MSH « Quelles sciences humaines et sociales pour le 21e siècle ? », organisée par la MRSH de Caen avec la collaboration du RNMSH. Ce fut l'occasion de rencontrer des collègues d'autres MSH mais aussi des représentants du Ministère de l'Enseignement Supérieur et de la Recherche.
- Fiche PUDC sur le site RNMSH disponible à l'adresse suivante : http://www.msh-reseau.fr/IMG/pdf/plateforme_universitaire_donnees_caen.pdf (voir en annexe)

CEREQ et réseau des centres associés

Compte tenu de la trajectoire des membres de la PUDC (Cyril Coinaud était chargé d'étude du CAR Céreq et Clotilde Lemarchant co-directrice), des liens privilégiés demeurent noués avec le Centre associé au Cereq de Caen (Centre d'études et de recherche sur les qualifications), qui dynamisent et orientent certaines activités de la PUDC, localement (participation à une ANR InEduc, voir ci-dessous) et en lien avec la « maison-mère » de Marseille. Le réseau des 17 centres associés au Céreq en France est actuellement en pleine restructuration. Récemment Cyril Coinaud et Clotilde Lemarchant ont participé (en vidéoconférence) à un « Atelier d'ouverture » du Céreq, Marseille, le 4 avril 2013 : « Données en ligne, communautés de chercheurs, la révolution en marche des SHS, où intervenait Roxane Silberman.

Adossement à l'enseignement supérieur

Dans le cadre des formations aux méthodes quantitatives réalisées à l'université de Caen en 2011-2012 et 2012-2013, Cyril Coinaud a participé à l'enseignement des « méthodes quantitatives » pour les étudiants de Licence 1 de Sociologie ainsi que pour l'année universitaire en cours à l'enseignement de la valeur « Informatique et bases de données » pour le parcours Management de l'Economie Sociale et Solidaire du master Management du Social et de la Santé 1^{ère} année de l'Institut d'Administration des Entreprises de l'université de Caen.

De plus, dans le cadre de la formation au sein de la PUDC, une formation au logiciel SAS a été prodiguée du 30 mars 2012 au 4 mai 2012 à 3 doctorants (2 en économie et 1 en psychologie) durant 10 demi-journées.

Adossement à la recherche

La PUDC est impliquée dans le projet **ANR InEduc** portée par le laboratoire ESO. Ce projet vise à identifier les inégalités éducatives liées aux contextes et espaces de vie des jeunes âgés de 11 à 15 ans (période scolaire du collège) en France. L'objectif est d'analyser différents contextes qui influencent les parcours éducatifs des jeunes : celui de l'institution scolaire, celui des loisirs sportifs et culturels et celui, transversal aux deux précédents, des environnements numériques. La dimension comparative de ce projet, qui traitera, sur le plan quantitatif, de trois territoires régionaux (Aquitaine, Basse-Normandie et Bretagne) et sur le plan qualitatif, d'espaces locaux sélectionnés dans ces régions, vise à enrichir la connaissance de la diversité des politiques d'éducation mises en œuvre et plus globalement des types d'effets de lieu sur les parcours éducatifs des jeunes. La dimension longitudinale de ces parcours sera explorée diversement selon les trois types de contextes éducatifs choisis.

La PUDC participe à la partie quantitative de la recherche (dirigée par Patrice Caro du laboratoire ESO-Caen et directeur du CAR Céreq de Caen). Elle consiste en la passation d'environ 3500 questionnaires dans les collèges des 3 régions précitées ainsi que la passation auprès des parents de ces mêmes collégiens. L'implication de la plateforme se situe au niveau de la mise en œuvre du protocole d'enquête :

- Elaboration et maquettage des 2 questionnaires.
- Nettoyage et codage des questionnaires après passation.
- Saisie automatisée des questionnaires (via le logiciel Data Scan).
- Apurement des bases de données obtenues.

Adossement à des missions internes à l'université

Créer un Observatoire « genre, égalité et mixité à l'UCBN » (En lien avec la charge de mission égalité femmes-hommes, et avec la CPED, conférence permanente des chargé-e-s de mission égalité, diversité)

Contexte

En avril 2011, la Présidente de l'université Josette Travert a demandé à Clotilde Lemarchant de devenir co-référente égalité hommes-femmes à l'université de Caen avec Patrice Georget et, compte tenu de ses travaux de recherche sur le genre et l'orientation scolaire, de constituer un groupe d'observation « genre, trajectoires de formation et ressources humaines à l'université de Caen Basse-

Normandie ». Elle confie ce dossier à M. Pierre Sineux, alors Vice-président de l'université (CEVU), dossier engageant les co-référents égalité des chances à l'UCBN, la PUDC, le service inscriptions DEVE ; le service ressources humaines, le service de la base Apogée au CRISI, Espace Orientation-Insertion; la cellule d'aide au pilotage ; le correspondant Informatique et Liberté.

Il s'agit de porter un regard sur les réalités sociales de notre établissement en fonction d'une problématique de genre et d'intégrer davantage ces réalités au projet d'établissement et au bilan social. Les réflexions en termes d'égalité professionnelle sont déclinées en deux domaines de recherche et d'action :

- **Ressources humaines** : à cet égard, un premier travail a été effectué en 2011 avec le mémoire de Lucie Hémary, master 2 psychologie sociale, encadré par Patrice Georget, MCF. Les deux co-référents égalité hommes-femmes restent en lien avec la DRH afin de définir de nouveaux thèmes privilégiés d'observation (situations et trajectoires des BIATOSS, etc).

- **Formation** : il s'agit d'observer les trajectoires des étudiant-e-s en privilégiant la grille de lecture du genre (orientation vers les diverses spécialités des filles et des garçons, rythme des formations). Un premier travail a été réalisé par la PUDC, en lien avec l'observatoire de l'université, et le Céreq. Le CEVU s'est dit très intéressé par un tableau de bord sur des données sexuées et interprétées. Faire un état des lieux annuel sur les formations servira de levier éventuel pour le CA de l'UCBN. Ce projet d'étude a été inscrit dans le contrat d'établissement.

Rapport « genre et parcours de formation à l'UCBN »

Un rapport sur les parcours de formation à l'UCBN dans une perspective de genre est presque finalisé et sera remis prochainement au président de l'université, Pierre Sineux. Y figure une liste de formations détaillées, leurs effectifs en fonction de la variable sexe, afin d'interroger la mixité. Un autre travail est davantage ciblé sur une étude fine des parcours selon le sexe des étudiant-e-s. La finalité du traitement est de permettre de comparer les flux (entrées, sorties, réorientation) des étudiants de l'université de Caen primo-entrants et néo-bacheliers à la rentrée universitaire de 2007 dans les différentes filières d'enseignements et selon le genre. Les données concernent les primo-inscrits néo-bacheliers de l'Université de Caen à la rentrée de l'année universitaire 2007-2008 (extraction de la base APOGEE).

Cette tâche, aux confins de la mission égalité femmes-hommes, de la présence active de la PUDC et des compétences en matière de recherche en études de genre, pourra être pérennisée et développée en fonction de la politique du genre que se donne l'établissement. Ce thème est en développement actuellement dans un contexte national volontariste sur le sujet : un GIS « Institut du genre » a été créé en 2012. La CPED (conférence permanente des chargé-e-s de mission égalité, diversité), fondée en 2011, relaie au niveau national ces diverses expérimentations en cours de réalisation dans les universités françaises.

Clotilde Lemarchant a reçu une décharge de cours de 20 heures en 2012-2013 pour ce faire. Il serait opportun de prendre en considération les besoins d'un poste d'ingénieur-e à l'avenir ciblé sur ce sujet.

Projets, perspectives et questions

Un contexte européen en pleine effervescence

Ces dernières années ont été le témoin de changements importants en matière de données SHS, tant sur le plan des statuts, de la production de données, que de l'organisation et de la gouvernance des institutions qui en ont la charge. Sur la feuille de route européenne, les très grandes infrastructures de recherche prennent en compte les SHS. Le CESSDA (Council of European Social Science Data Archives) et un consortium ERIC (European Research Infrastructure Consortium) ont été mis en place.

Des outils qui se structurent, des statuts qui se précisent au niveau national

Aujourd'hui, le cadre institutionnel national est structuré autour de deux TGIR : PROGEDO (production et gestion des données) et ADONIS- Corpus SHS (relatif au documentaire, à la bibliothèque numérique scientifique). Les plates-formes de données relèvent de PROGEDO, qui rassemble d'une part le réseau Quetelet, d'autre part, des actions de soutien à la production de données pluridisciplinaires (notamment ESS –european sociological survey- et SHARE – sur le vieillissement) pour lesquelles la France cherche des financements de grandes cohortes).

Le CASD, Centre d'accès sécurisé à distance, devient massivement le passage obligé : le Ministère de l'agriculture a passé une convention et d'autres sont en cours (DARES, etc). Cette nouvelle organisation permet que des données agrégées soient en ligne, disponibles et accessibles (cf. procédure NESSTAR et les métadonnées). Les procédures auprès du comité du secret sont en cours de simplification. Les plates-formes universitaires de données permettent la mutualisation des SD box permettant de recevoir à distance les fichiers demandés et sécurisés.

Une formation de correspondant informatique et liberté est envisagée en 2013 organisée par la CNIL.

De nouvelles questions : gratuité, confidentialité

Sous la pression croissante des acteurs économiques pour accéder à ces données multiples, la question de la définition du périmètre des utilisateurs se posent sous un nouveau jour. A qui ouvre-t-on ces données (instituts de sondage ? écoles de management ?). Se pose pareillement la question de la gratuité des données et celle de leur confidentialité.

Des enjeux localement : Les plates-formes ont pour vocation de promouvoir l'utilisation des données grâce à une implantation locale, au plus près des utilisateurs, au carrefour des laboratoires de recherche et de l'enseignement supérieur. Le fait que ces données se multiplient, de sources différentes et sous des formes variées, constitue un atout régional.

Pour l'avenir, la pérennité de la PUDC repose sur une condition : assurer le poste de l'ingénieur statisticien et les frais de fonctionnement de la plate-forme.

Bilan Financier

Budget de fonctionnement

Les frais de la PUDC reposent en premier lieu sur des charges de personnel, le montant d'un salaire annuel à temps plein d'un ingénieur d'étude statisticien confirmé : 29 767,56 €.

Le tableau ci-dessous présente les dépenses liées aux autres frais de fonctionnement de la PUDC.

Déplacements	400,67 €
Licences et logiciels	910,63 €
Frais reprographie	215,29 €
Frais téléphoniques	250,00 €
Frais affranchissement	60,00 €
Frais consommables informatiques	160,00 €
TOTAL DEPENSES 2012	1 996,59 €

Bilan des dépenses de fonctionnement de la PUDC

Budget prévisionnel

Les prévisions pour le budget de fonctionnement de 2013 sont plus élevées : il faudra prévoir le coût d'un logiciel supplémentaire, data scan d'environ 3000 €. Les coûts de déplacements seront aussi plus élevés puisque nous envisageons de participer à la conférence annuelle IASSIST sur les données internationales (environ 1 400 €).

Annexes

Compte-rendu de la réunion du comité de pilotage 2012

Compte-rendu du Comité de pilotage n°1 de la plate-forme universitaire de données de Caen (PUDC)

Le 3 avril 2012, Université de Caen UCBN, MRSH, salle 028

Présent-e-s :

Roxane Silberman, secrétaire générale du Comité de Concertation pour les Données en SHS, Jean-Louis Lagarde, Vice-président de l'université de Caen, Christophe Leconte, direction de la recherche, Magali Tencé, représentant le CNRS en Région, Pascal Buléon, directeur de la MRSH, Annie Laurent, responsable administrative de la MRSH, Daniel Delahaye, représentant les écoles doctorales, Vincent Merlin (CREM), Jean-Louis Lenhof (CRHQ), Joël Brée (NIMEC), directeurs de laboratoires de recherches en SHS de Basse-Normandie, Anne Pellissier, représentant les sciences de l'éducation et le CERSE, Clotilde Lemarchant et Cyril Coinaud, la responsable et l'ingénieur de la plate-forme.

Excusé-e-s :

Patrice Caro (ESO/ Céreq), Philippe Madeline (ESO), Christian Papilloud, Charlotte Le Van (département de sociologie /Cereve), Thierry de Saint-Gérand (geosyscom).

Pascal Buléon ouvre la séance et remercie les participants de leur présence en cette journée très remplie (élections des membres des conseils à l'UCBN).

Clotilde Lemarchant remercie les participant-e-s et commente le rapport d'activité 2011 de la plate-forme universitaire de données distribué à chacun-e : créée fin janvier 2011, et hébergée au cœur de la MRSH, la PUDC connaît au bout d'un an une certaine dynamique (voir rapport d'activité 2011 joint au compte-rendu) : 56 utilisateurs et utilisatrices différent-e-s ont contacté et rencontré à plusieurs reprises Cyril Coinaud, l'ingénieur : en premier lieu des doctorant-e-s et, en second lieu, des chercheur-e-s, pour des séances de formation SAS, pour un soutien technique d'aide à la création d'enquête par questionnaires, mais aussi d'accès à de grandes enquêtes sur leur sujet de recherche,... La plate-forme est aussi associée à des projets collectifs de recherche (Labex, ANR). Elle est adossée à un travail de réflexion sur le genre et les trajectoires des étudiant-e-s et des personnels de l'université de Caen, en lien avec la DEVE, la DRH et l'observatoire de l'université. Clotilde Lemarchant rappelle qu'elle a été nommée « co-référente égalité diversité » avec Patrice Georget, en 2011 par Mme Travert.

Pascal Buléon dit sa satisfaction devant la dynamique du fonctionnement de la plate-forme de données et remarque son utilisation par l'institution universitaire. Il rappelle que l'enveloppe budgétaire très attendue par l'université est arrivée fin mars. Il s'adresse à Jean-Louis Lagarde, vice-président et l'interroge sur les projets d'attributions budgétaires : il n'est pas sûr qu'il n'y ait pas de ligne budgétaire attribuée à la PUDC...

Jean-Louis Lagarde commente rapidement ce budget de l'université, qui ne connaît pas vraiment d'augmentation. Il sera difficile de revendiquer 400 000 euros mais possible de financer un ingénieur et de poursuivre le travail en cours. L'UCBN est heureuse de voir la plate-forme fonctionner et s'engage à en assurer son financement. Il faudra voir comment l'intégrer dans une sorte de grande plate-forme SHS.

Vincent Merlin note l'appétence des doctorants du CREM envers la PUDC, intérêt qui pourrait aller croissant compte tenu de l'inclusion du CREM dans une fédération de recherche et le développement récent de la recherche appliquée, notamment sur les thèmes du travail et de l'emploi. Il présente le module de formation à l'économétrie et à l'analyse de données accessible aux étudiants rattachés au CREM ainsi qu'au NIMEC (avec un module plus qualitatif et un autre d'économétrie « hard »). Il faudrait penser, selon lui, à intégrer un module de ce type dans le plan de formation des écoles doctorales. Une réflexion est à mener en termes de collègues actuellement.

Joël Brée rappelle qu'une séance de présentation de la PUDC au NIMEC a eu lieu il y a un mois, ce qui a permis de comprendre le fonctionnement concret de ce service. Le besoin premier des doctorants à l'IAE réside dans la formation aux statistiques. La trop faible connaissance en la matière crée des inhibitions, de l'autocensure. Les enquêtes « Valeurs » suscitent pourtant tout particulièrement leur intérêt. Joël Brée met lui aussi l'accent sur l'intérêt d'une formation dans le cadre des écoles doctorales.

Daniel Delahaye propose de relayer les services de la PUDC auprès de l'école doctorale, et de les faire apparaître dans le programme des écoles doctorales, en particulier auprès de l'école doctorale spécialisée en SHS. Il reste à en chiffrer le coût.

Anne Pellissier souligne elle aussi le besoin de formation et d'information sur la culture des grandes enquêtes en sciences de l'éducation. Intervenant en médecine, elle rappelle les liens entre sciences humaines et cette discipline et propose d'informer régulièrement les doctorants en médecine des services de la PUDC.

Christophe Leconte rappelle les travaux de Guy Benoit aux confins de la médecine et des SHS, à l'UCBN.

Magali Tencé dit que le CNRS se tient à disposition de la PUDC pour la faire connaître dans les réseaux du CNRS.

Roxane Silberman insiste sur le fait qu'un dispositif national ne peut fonctionner que s'il existe un tissu local et se dit satisfaite de voir la dynamique avancer. La diversité des usagers de Caen est une très bonne chose. A l'avenir, d'autres plates-formes devraient être créées en France, à Marseille, à Toulouse, chacune pouvant avoir sa spécificité. Le CASD (centre d'accès sécurisé à distance) voit son rôle augmenter. La dimension de l'archivage est amenée à connaître un essor, permettant la récupération d'enquêtes locales. En outre, des liens se développent entre tissu local et tissu

international. Face à cette effervescence, la question qui se pose à Caen est la pérennité du poste de Cyril Coinaud.

Jean-Louis Lagarde se dit confiant pour l'avenir de la plate-forme et affirme que l'université soutiendra et le poste d'ingénieur et la plate-forme. A terme, en cas de développement de l'activité de la PUDC, si un second poste d'ingénieur d'étude est nécessaire, les laboratoires et services devront alors participer financièrement. Mais le noyau dur doit être assuré par l'établissement.

Par Clotilde Lemarchant, MCF de sociologie, responsable de la PUDC.

FICHE PROFIL DE POSTE

BAP : D

CORPS : Ingénieur-e d'études

NATURE : Contractuel

ETABLISSEMENT : Université de Caen Basse-Normandie

UFR ou SERVICE : Maison de la Recherche en Sciences Humaines, plate-forme universitaire de données de Caen (PUDC)

DESCRIPTION DU POSTE

Contribuer à la diffusion, auprès de chercheurs et d'étudiants normands, de données d'enquêtes et de bases de données dans le domaine des sciences humaines. Aider les chercheurs dans la recherche et l'utilisation de ces données.

ACTIVITES

Accompagner des chercheurs et étudiants dans leur recherche et acquisition de fichiers issus de grandes enquêtes : choix des fichiers adaptés et des organismes, déontologie à respecter dans l'utilisation de ces fichiers, installation des fichiers sur les postes de travail, démarches administratives, mise en œuvre de la documentation de travail (questionnaire, plan de sondage, liste et contenu des variables...), aide à l'interprétation des analyses statistiques (analyse de variance, classification, régressions, chi2...)

Organiser des formations en matière de statistique (interprétation des tableaux de contingence, description d'une distribution statistique, tests, méthodes d'analyse ou d'exploration des données) et d'informatique (gestion de fichiers, navigation sur le web, logiciels statistiques, notamment SAS).

Poursuivre le travail de communication au sujet de l'existence de cette plateforme déjà entrepris depuis un an auprès des différents laboratoires de sciences humaines et unités de formations de l'université de Caen UCBN.

Participer à la création et au fonctionnement d'un groupe de travail « genre, formations, professions au sein de l'université », en partenariat avec l'observatoire de l'université, le SUOIP, le service inscriptions DEVE, la cellule d'aide au pilotage. Soutenir la phase de traitement de données de l'ANR INEDUC (2012-2015) en collaboration avec le laboratoire de géographie sociale ESO.

Suivre les évolutions des programmes d'enquêtes de la statistique publique et veiller au calendrier de leur diffusion.

COMPETENCES ET CONNAISSANCES SOUHAITEES

Maîtrise avancée des procédures de gestion et d'analyse statistique d'un logiciel d'analyse de données, SAS de préférence. Des connaissances d'un autre logiciel seraient appréciées.

Maîtrise du système d'exploitation Windows, des connaissances de Linux seraient appréciées.

Maîtrise des outils bureautiques usuels.

Bonne culture en sciences humaines et sciences sociales (sociologie, économie, histoire, géographie...) et notamment en matière de production de données d'enquêtes par sondage et de base de données.

Connaissance approfondie des normes internationales et nationales en matière de nomenclatures usuelles dans les SHS (emploi, éducation, géographiques...).

Connaissance approfondie des règles déontologiques en matière de grandes enquêtes.

Maîtrise du traitement de données statistiques en SHS.

Contribuer à la réflexion sur les techniques et méthodes en Sciences Humaines et Sociales

Aisance en anglais, au moins écrit.

QUALITES REQUISES

Aptitude au travail en équipe dans un contexte de projet.

CONTEXTE : La plate-forme universitaire des données de Caen est partenaire du réseau Quételet. Elle est amenée à nouer des relations privilégiées avec l'ADISP (Archives de données issues de la statistique publique) du Centre Maurice Halbwachs, le service des enquêtes de l'INED et le CDSP (Centre de données socio-politiques) de Sciences Po Paris. Enfin, elle est amenée à échanger avec les deux autres plates-formes existantes de Lille et de Lyon.

FICHE PROFIL DE POSTE

BAP : D

CORPS : Ingénieur-e d'études

NATURE : Contractuel

ETABLISSEMENT : Université de Caen Basse-Normandie

UFR ou SERVICE : Maison de la Recherche en Sciences Humaines, service Pole égalité femmes-hommes

DESCRIPTION DU POSTE

Produire régulièrement un état des lieux d'une part des trajectoires des étudiant-e-s de l'université de Caen par sexe et par filière de formation, d'autre part des caractéristiques et évolutions de carrières des personnels de cette université (affectation, évolution de carrière, formation continue, responsabilités,...) dans une perspective de genre.

Construire des données chiffrées et réfléchir aux indicateurs statistiques les plus pertinents sur ces questions. Une attention sera portée sur la dimension longitudinale et les évolutions depuis 15 ans.

S'informer des états des lieux semblables éventuellement réalisés dans les autres universités, en lien avec les membres de la CPED (conférence permanente des référents égalité diversité).

ACTIVITES

Repérage, recueil, mise en ordre et synthèse des données disponibles dans les différents services de l'université ayant trait aux étudiants et aux personnels. Ces activités se feront en lien avec différents services de l'université, notamment la DRH et le service inscriptions des étudiants DEVE.

Construction de nouvelles données sur la base d'enquêtes complémentaires centrées sur le genre.

Diffusion et exploitation régulière de l'information dans ce domaine auprès du président de l'université.

Etablissement d'un bilan de la situation des hommes et des femmes à l'université à partir des données disponibles du bilan social.

Proposition de pistes de travail et d'un plan d'action en lien avec les co-référent-e-s pour « l'égalité des droits entre les femmes et les hommes » de l'université de Caen.

Intégration de la problématique du genre dans les formations et les conditions d'accès au monde du travail, en collaboration avec le Vice-président du CEVU.

COMPETENCES ET CONNAISSANCES SOUHAITEES

Maîtrise avancée des procédures de gestion et d'analyse statistique d'un logiciel d'analyse de données, SAS ou/et R.

Maîtrise du système d'exploitation Windows, des connaissances de Linux seraient appréciées.

Maîtrise du traitement de données statistiques en SHS.

Maîtrise des outils bureautiques usuels.

Bonne culture en sciences humaines et sciences sociales (sociologie, économie, histoire, géographie...) et en matière de production de données d'enquêtes par sondage et de base de données. Intérêt pour la question du genre.

QUALITES REQUISES

Aptitude au travail en équipe dans un contexte de projet.

CONTEXTE

Le service Pole égalité femmes-hommes est amené à travailler en collaboration avec la PUDC (plateforme universitaire des données de Caen) et avec le soutien de différents services de l'université : le service inscriptions DEVE, le service de la base Apogée au CRISI, la cellule d'aide au pilotage, la Direction de la Communication, la Direction des Ressources Humaines et en lien avec l'Observatoire de l'université.



Nom et prénom des responsables : Clotilde LEMARCHANT

Téléphone : +33 (0)2 31 56 62 99

Mail : clotilde.lemarchant@unicaen.fr

Cyril COINAUD

Téléphone : +33 (0)2 31 56 61 43

Mail : cyril.coinaud@unicaen.fr

Objectifs

Une plateforme universitaire de données vise à faciliter l'accès aux grandes enquêtes et données statistiques intéressant les sciences humaines et leur utilisation.

Outil de proximité à l'égard des étudiants, enseignants et chercheurs de l'université de Caen Basse-Normandie dans toutes les disciplines des sciences humaines, la PUDC offre ses services dans la recherche de fichiers adaptés aux recherches, les démarches administratives avec le réseau Quetelet et la récupération et l'installation des bases de données. Elle propose un accompagnement méthodologique et des formations adaptées, de la conception des questionnaires jusqu'au traitement statistique.

Equipements

La MRS H de Caen a mis à disposition deux bureaux pour les membres de la PUDC et pour recevoir les étudiants et chercheurs utilisateurs des services de la plateforme.

La PUDC a fait l'acquisition de logiciels d'analyse de données : SAS, SPSS, STATA, R ainsi que d'un logiciel de recueil de données d'enquêtes par questionnaires (DATA-SCAN). Elle est aussi un lieu de mise à disposition de livres et documentations d'enquêtes et de méthodes statistiques.

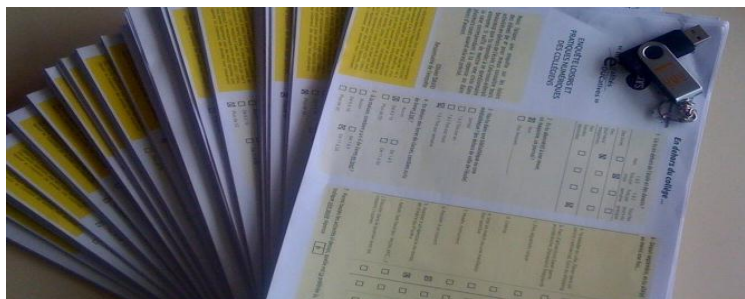
Domaines d'activité

La PUDC permet de connaître les organismes, les grandes enquêtes les plus appropriées à une recherche, les démarches pour obtenir les fichiers, la déontologie à respecter dans l'utilisation de ces fichiers. Elle propose un accompagnement personnalisé en traitement de données et analyse statistique : interprétation, description d'une distribution, tests statistiques, méthodes d'analyse, ainsi que l'utilisation des logiciels statistiques.

Elle a enfin pour fonction d'archiver de façon pérenne les données des recherches réalisées dans le cadre de l'université de Caen, quantitatives ou qualitatives.

Partenariat

La PUDC est partenaire du GIS Réseau Quetelet qui rassemble des producteurs et diffuseurs de données de la statistique publique : INED, Sciences-Po, CMH-Adisp. Elle travaille en réseau avec d'autres plateformes universitaires de données.



Site internet : <http://www.unicaen.fr/recherche/mrsh/pudc>