



plateforme universitaire
de données de Caen

La PUDC, plateforme universitaire de données à Caen

2010-2012

*Maison de la Recherche en Sciences Humaines
CNRS - Université de Caen Basse-Normandie*

RESPONSABLE
CLOTILDE LEMARCHANT

M R S H | NORMANDIE - CAEN
Maison de la Recherche en Sciences Humaines
CNRS - UNIVERSITÉ DE CAEN



plateforme universitaire
de données de Caen

Rapport d'activité de la PUDC 2011

(Bilan au 12 mars 2012)

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

I- Création d'une plateforme universitaire de données à Caen (PUDC) en janvier 2011

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

1- Un outil en réseau

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.

La PUDC est financée dans le cadre d'un PPF du précédent quadriennal qui prend fin en décembre 2011. 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.

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.

La PUDC travaille en partenariat avec les deux autres plateformes. Pour le montage du projet, les rencontres avec la plateforme lilloise ont été déterminantes et fructueuses, source de collecte de méthodes, bonnes pratiques, témoignages de caractéristiques du bon fonctionnement d'une plateforme de diffusion de données au sein d'une université : les aspects techniques (matériel nécessaire, locaux, outils informatiques...) les moyens les plus efficaces de se faire connaître, les aspects saisonniers du travail, l'intérêt d'ouvrir l'accès de la plateforme aux étudiants de master2, etc.

Un comité de suivi pluridisciplinaire a été mis en place afin d'accompagner localement le lancement de la PUDC qui s'est réuni le 11 avril 2011 (voir en annexes). La présentation d'un rapport d'activité intermédiaire a eu lieu le 25 novembre 2011 (voir compte rendu en annexes). Un comité de pilotage, plus formel, se réunira pour la première fois le 3 avril 2012. Il se compose d'une dizaine 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/la responsable de la MRSH
- Les représentants des laboratoires de recherches en SHS de Basse-Normandie
- Les représentants des responsables de masters
- Les représentants des écoles doctorales
- Le/la responsable de la plateforme

2- Un outil de la MRSH Normandie Caen

Hébergée à la MRSH, 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 en formation continue à l'UCBN.

Le financement du poste a été reconduit pour 6 mois jusqu'au 30 juin 2012.

Elle est localisée à la Maison de la recherche en sciences humaines (MRSB) 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 MRSB a soutenu la PUDC en lui procurant des locaux et finançant ses premiers frais, entre autres les deux postes d'ordinateurs nécessaires à son fonctionnement.

La plateforme de données est accueillie et portée par la MRSB 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, « grands » ou modestes. La plateforme de données est une de ces ressources qu'il faudra, à l'avenir, mutualiser.

En effet, le CASD (Centre d'accès sécurisé distant) confirme l'incitation à développer un réseau national de plateformes universitaires. Des projets sont en cours de discussion dans l'ouest de la France, à Toulouse notamment. Lors d'une réunion du GIS réseau Quételet du 14 avril 2011, le CASD, représenté par Kamel Gadouche, propose que le CASD et les plateformes de données se réunissent régulièrement afin d'harmoniser les actions et anticiper les exigences européennes futures en matière de données sensibles et sécurisées. En effet, à l'avenir, il est probable que les chercheurs disposent moins facilement des données, qui ne seront consultables que dans des locaux sécurisés (codes, badges, biométrie, caméras...). En ce cas, privilégier des endroits partagés (les MSH) plutôt que les laboratoires serait opportun. Plus largement, les normes européennes vont être amenées à s'harmoniser ce qui va nécessiter un accompagnement plus grand des chercheurs.

Lors de la visite des représentants de l'AERES à la MRSB de Caen, le 16 juin 2011, la plateforme des données a été un outil remarqué.

Elle a aussi été présentée lors du Conseil de gestion de la MRSB le 6 juillet 2011.

II- Communiquer, s'informer, se former

1- La communication au sujet de la plateforme

a. Mise en forme de la page sur le site de la MRSB

La MRSB dispose d'un site internet et d'un outil de diffusion efficace mis à la disposition de la PUDC. Un site internet est disponible à l'adresse suivante : <http://www.unicaen.fr/recherche/mrsh/pudc>

b. Les entrevues avec les responsables d'UFR et de laboratoire

Au cours de l'année 2011, tout un travail de présentation des services de la plateforme a été réalisé auprès des collègues et étudiants de l'UCBN. Notamment, des séances de présentation ont été faites lors de réunions de laboratoire, de département ou d'UFR, ou lors de cours de méthodologie.

Ont été rencontrés :

- Le CREM (économistes)
- GEOPHEN (géographie physique)
- ESO (géographie sociale)
- L'UFR Sciences de l'éducation
- Le département de sociologie
- L'IUP MSS (politiques sociales et santé publique)
- l'UFR STAPS.
- le NIMEC (IAE).

Des liens ont également été tissés avec le CRHQ (histoire quantitative), la psychologie sociale, le CRAHAM (histoire médiévale), les MASS (Mathématiques appliquées).

c. La plaquette d'information

Une plaquette d'information a été réalisée par le service communication et infographie de la MRSH (voir document ci-joint en annexe). Une affiche est en cours de réalisation.

d. e-phenix

La PUDC a été présentée au sein de la publication de l'université de Caen-Basse-Normandie, le e-phenix n° 44.

La plate-forme universitaire des données de Caen - PUDC

La Plate-forme universitaire des données de l'université de Caen (PUDC) 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.

Cette plate-forme est localisée à la maison de la recherche en sciences humaines (MRSH), bureau n° 109 (1^{er} étage).

www.unicaen.fr/recherche/mrsh/pudc
Contact : pudc@unicaen.fr

Extrait du e-phenix n°44.

e. La lettre de la MRSH

La PUDC a été présentée en première page du numéro 120 de la lettre d'information de la MRSH en janvier 2012.

MRSH **Lettre d'info**
NORMANDIE - CAEN
Maison de la Recherche en Sciences Humaines
CNRS - UNIVERSITÉ DE CAEN

NUMÉRO 120 - JANVIER 2012

EN LIGNE CE MOIS-CI
sur <http://www.unicaen.fr/recherche/mrsh/forge>

La forge numérique

À écouter

- *La révolution napoléonienne mise en images*
ALAIN HUGON
- *Violences et conflits dans la société rurale bretonne d'Ancien Régime: l'apport croisé des archives judiciaires et des chansons de tradition orale*
EVA GUILLOREL
- *Les modalités de la conflictualité paysagère*
HERVÉ DAVODEAU
- *Conformité sociale et normes scolaires*
ALAIN VERGNIoux

Alain Vergnioux est agrégé de philosophie, docteur en philosophie, professeur émérite de l'université de Caen où il a enseigné au département des Sciences de l'éducation et a été rédacteur en chef de la revue *Le Télémaque* (philosophie de l'éducation) ainsi que de la revue *Les sciences de l'éducation. Pour l'ère Nouvelle* (publication du CERSE).

- *La violence au quotidien des enfants des rues en Bolivie (La Paz) ou le concept d'enfant dans une autre culture*
CHARLES-ÉDOUARD DE SUREMAIN
- *La valeur « conformité » à l'épreuve des collégiens et lycéens*
MONIQUE WACH, MARIE-HÉLÈNE GROS
- *Se vêtir au collège : les violences insidieuses de la conformité*
NATHALIE DUPONT, CAROLINE GIRARD
- *Le sens du temps libre dans l'emploi du temps de jeunes collégiens*
JULIE DELALANDE, SYLVIE JOUENNE

À voir

- *Les missions du Conservatoire du littoral*
RÉGIS LEYMARIE

La photo du mois

Le cochon est joyeux, toujours le jour !
Gilles-Paul Robélot

La Maison de la Recherche en Sciences Humaines vous souhaite une bonne année 2012

Plate-forme Universitaire de Données de Caen (PUDC)

La Plate-forme Universitaire de Données de Caen est un espace dédié aux données quantitatives et installé à la MRSH depuis janvier 2011. Vous y trouvez un accès aux données issues de la statistique publique et des grandes enquêtes nationales ou internationales, ainsi qu'un accompagnement méthodologique et technique favorisant leur utilisation.
Responsable scientifique : Clotilde Lemarchant

La PUDC est un service proposé par la MRSH Normandie Caen en direction des étudiants, doctorants et chercheurs de sciences humaines et sociales. Elle a la double vocation de faciliter l'accès gratuitement (uniquement dans le cadre de la recherche) à des données statistiques diffusées par le réseau Quetelet et d'assister individuellement tout utilisateur à chaque étape de l'élaboration d'une recherche quantitative et d'une analyse statistique (conception d'un questionnaire, recherche de données adaptées, utilisation de logiciels de traitement statistique, description d'une distribution, tests statistiques, exploitation des résultats, etc.). Elle vous propose également une assistance à la production de vos propres données d'enquêtes.

Travaillant en réseau avec les autres plates-formes de données et les partenaires du réseau Quetelet, la PUDC facilite l'accès aux enquêtes de grands fournisseurs de données (INSEE, services statistiques de ministères (Éducation Nationale, Santé,...) et à la documentation associée à ces données.

La PUDC est disposée également à apporter des réponses collectives pour des formations complémentaires aux méthodes quantitatives ainsi qu'aux logiciels d'exploitation statistique (SAS, R, SPSS, STATA, EXCEL...).

L'équipe, composée de Clotilde Lemarchant (MCF de sociologie) et Cyril Coinaud (ingénieur statisticien), se tient à votre disposition pour toute information supplémentaire, au bureau SH 109 de la MRSH (tél. 02 31 56 61 43 ; cyril.coinaud@unicaen.fr)

2- Les séminaires, les colloques et la veille d'information

Etre en mesure de proposer les enquêtes et les fichiers adéquats nécessite d'être toujours en veille avec les outils et les méthodes en matière de production et diffusion de données, à un niveau national et international. Cyril Coinaud a assisté à quelques événements en la matière.

a. IASSIST

IASSIST est une organisation internationale des professionnels de l'information et des données numériques pour la recherche et l'enseignement dans les sciences sociales qui organise un séminaire annuel de rassemblement international.

Durant le temps de ce séminaire (du 30 mai au 3 juin 2011 à Vancouver et organisé par la bibliothèque universitaire Simon Fraser (SFU) et par la bibliothèque de l'université de Colombie Britannique (UBC)), les concepts ont été développés autour de 3 grands axes qui peuvent être vus d'un point de vue chronologique dans le cycle de vie d'une enquête : résoudre les problèmes liés à la qualité pour la production des données, les standards pour l'archivage (NESSTAR, DDI,...) et les moyens logiciels de diffusion et le développement d'outils d'analyse de données on-line.

Afin d'améliorer les données produites, il convient de s'intéresser aux moyens de leur production. Ainsi certaines interventions du séminaire portaient sur des sujets comme l'augmentation des taux de réponses par diverses méthodes (rémunération, cadeaux, ...), sur des systèmes de notations des données par les utilisateurs et sur un regard plus générale sur la qualité des données (« Verre à moitié vide ou à moitié plein ? »).

Une partie plus technique et opérationnelle a été largement appréhender tout au long de la conférence que cela lors d'atelier de travail (mise en pratique de la NESSTARisation d'une enquête) ou d'intervention sur différents logiciels pour la documentation des enquêtes (interfaces pour l'implémentation DDI des documents) produits « localement » par des centres d'archives selon leur problématique propre. D'autres encore ont présenté des modules de recherche de sites internet de leur centre d'archives ou encore levé certaines difficultés rencontrés comme l'appariement de différentes sources pour un même individu ou la documentation multi-langues.

Pour certains centres d'archives, l'heure n'est plus uniquement à la diffusion des données via leur site internet mais à l'exploration et à l'analyse des données. Ainsi, certains dotent leur site d'outils « statistiques » on-line et directement intégrés comme la cartographie, l'analyse de données longitudinales et de création d'extraction des données.

De manière moins formelle, des discussions autour des différences de structuration et historico-organisationnelles entre pays en ce qui concerne l'archivage et la diffusion de données, des particularismes des diverses enquêtes et recensements des populations selon les pays ont permis de m'alerter sur la complexité de comparaison et d'uniformisation des données internationales (absence de nomenclatures, modes de passation différents, ...) et ainsi de mieux comprendre les particularités du système français.

b. Le séminaire « Mesure des inégalités »

Le lundi 21 novembre 2011 à l'Institut Henri Poincaré, la Société Française de Statistiques et le Centre Maurice Halbwachs ont organisé ce séminaire sur la mesure des inégalités.

Résumé du séminaire :

Tout le monde a une idée de ce que sont les inégalités et en parle comme si c'était quelque chose de simple. Pourtant ce concept est complexe, intégrant d'emblée celui de comparaison: il s'agit en l'occurrence de comparer ce que possèdent les individus ou leurs probabilités d'accéder à un type de bien ou un statut, ce qui présuppose de pouvoir le mesurer de façon précise. Ces mesures-là ne sont déjà pas simples, notamment du fait du caractère multidimensionnel des richesses. Mais, même en imaginant que l'on sait les effectuer, la mesure des inégalités reste problématique: comment décider si une distribution est plus inégale qu'une autre? Le manque de définition formelle conduit à des résultats qui semblent paradoxaux. Au-delà de ces questions formelles le recueil de données robustes et fiables recèle des difficultés de plusieurs types: informations intimes, inconnues, inaccessibles, manipulées... Par ailleurs il existe plusieurs sortes d'inégalités: inégalités de revenus, de santé, d'éducation, de talent, d'emploi, environnementales... et, selon le critère de comparaison, inégalités générationnelles, de genre, géographiques, professionnelles ..., qui peuvent éventuellement s'entrecroiser; elles ne requièrent peut-être pas toutes la même méthodologie pour être calculées. Ces journées d'études invitent chercheurs et statisticiens à réfléchir à la façon dont on peut progresser vers une analyse réaliste, consensuelle et compréhensible.

c. La veille d'information.

La plateforme assure aussi une veille d'information à la fois sur les méthodes, les séminaires disponibles et les données. Son inscription dans des lettres d'information permet de repérer et relayer de l'information. La participation à des séminaires et des colloques à visée méthodologique s'inscrit aussi dans cette démarche.

d. L'inauguration de l'équipex Dime-SHS

Cyril Coinaud a assisté à l'inauguration de l'équipex Dime-SHS le 2 février 2012 à SciencesPo à Paris.

Ce projet porté par SciencesPo qui associe de nombreux partenaires (INED, Paris Descartes, GENES, GIS réseau Quetelet, EDF R&D, Telecom ParisTech) se compose de 3 sous-projets :

- DIME-SHS Web

- DIME-SHS Quali
- DIME-SHS Quanti

1) DIME-SHS Web.

Ce projet a pour ambition de prendre le web comme terrain. Pour cela, la création d'une interface mettra de circonscrire des corpus, de moissonner les contenus et d'utiliser des outils créés pour analyser le tout. Contrairement à d'autre « moteur de recherche de type Google », celui utilisé dans ce projet a pour ambition de privilégier la robustesse à la performance.

Ouverture du site en 2014.

Contact : mathieu.jacomy@sciences-po.fr

2) DIME-SHS Quali

Il a pour projet d'archiver et de diffuser des données qualitatives pour permettre des passerelles entre les infrastructures de recherche et les chercheurs.

Il permettra de mettre chapitrer les enregistrements et de manipuler les retranscriptions des entretiens. Il y a pour l'instant 3 enquêtes dans le catalogue.

Le travail de préparation pour la diffusion est très long. Il faudrait que les chercheurs anticipent un peu en élaborant leur matériau dans des langages compatibles avec la diffusion.

Contact : guillaume.garcia@sciences-po.fr

3) DIME-SHS Quanti

Cette partie s'occupe principalement de la production de données par 2 entrées : un panel internet et une plateforme téléphonique.

- Création de l'enquête ELIPSS : Enquête Longitudinale par Internet pour les Sciences Sociales. Sélection d'un panel de 6000 enquêtes (1500 pour le lancement) sur tablettes numériques avec connexion 3G distribuées gratuitement pour attraper les personnes qui n'ont pas internet chez eux avec en contrepartie, l'obligation de répondre tous les mois à un questionnaire d'environ 30 minutes. Les questionnaires auront un module permanent barométrique et un module changeant dans lequel pourront être posées les questions de chercheurs.
- Création d'une plateforme téléphonique pour centraliser la passation d'enquête par questionnaires téléphoniques.

Contacts : annesophie.cousteaux@sciences-po.fr

Stephane.legleye@ined.fr

3- Les formations suivies

Cyril Coinaud a suivi quelques formations complémentaires afin de cibler certaines compétences, notamment sur de nouveaux logiciels de traitement de données, adaptés aux sciences humaines et sociales.

a. Formation sur les petits échantillons

L'objectif de cette formation était, sur une petite série de données, de savoir choisir un test adapté aux questions que l'on se pose et aux données dont on dispose, de savoir comparer deux ou plusieurs échantillons, mesurer l'association entre plusieurs variables et de savoir mettre en œuvre ces tests et interpréter les sorties logicielles. Pour ce faire, la formation organisée s'est organisée sur trois jours consécutifs à partir du 5 juillet 2011.

b. Formation d'analyse de données sur le logiciel R

Afin de proposer une aide plus large que celle qui existe déjà à la plateforme, l'ingénieur statisticien a participé à une formation organisée par le CNRS pour le laboratoire « Centre Maurice Halbwachs » concernant l'analyse statistique sous le langage R.

L'objectif général de cette formation de trois jours était d'initier à la pratique sous R d'un certain nombre de techniques statistiques utilisées couramment en sociologie quantitative. L'accent est mis sur la pratique des techniques statistiques sur R.

La session a duré trois jours. Une journée et demie était consacrée à la découverte du logiciel, l'apprentissage de la syntaxe et de la structure orientée objet. Une journée et demi était consacrée aux principales techniques d'analyse de données à la française.

c. Achat de livres et matériel

L'achat de livres de méthodes statistiques appliquées aux sciences humaines et sociales s'inscrit à la fois dans le souhait de proposer une documentation adaptée aux étudiants et chercheurs et à la nécessité de rester en veille sur ces sujets en constante transformation.

III-Bilan des activités de la plateforme universitaire de données de Caen en 2011

La PUDC accueille déjà une cinquantaine de demandes de personnes différentes concernant les données et est sollicitée à la fois pour des formations et un travail d'archivage.

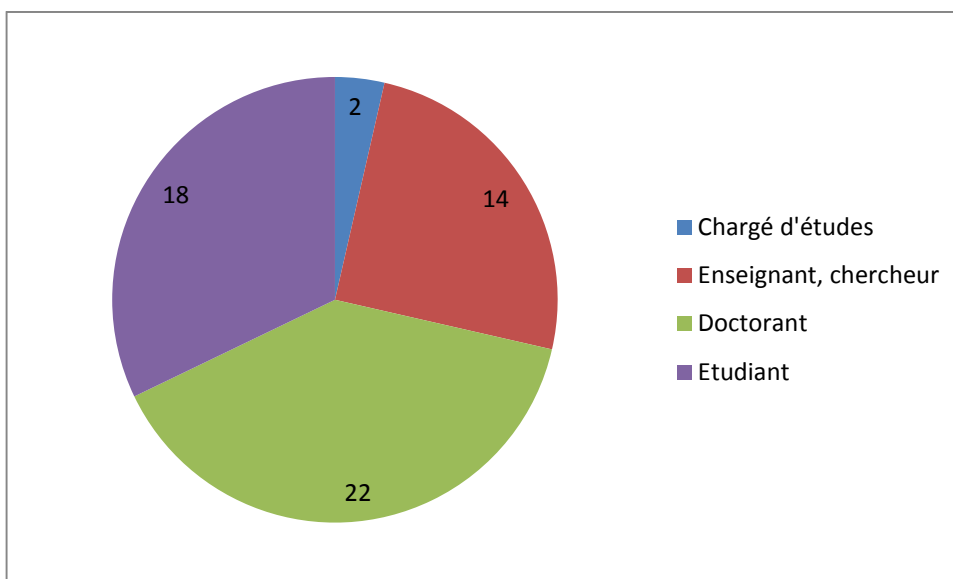
1-Le public qui fréquente la PUDC et les demandes exprimées

a. Le public

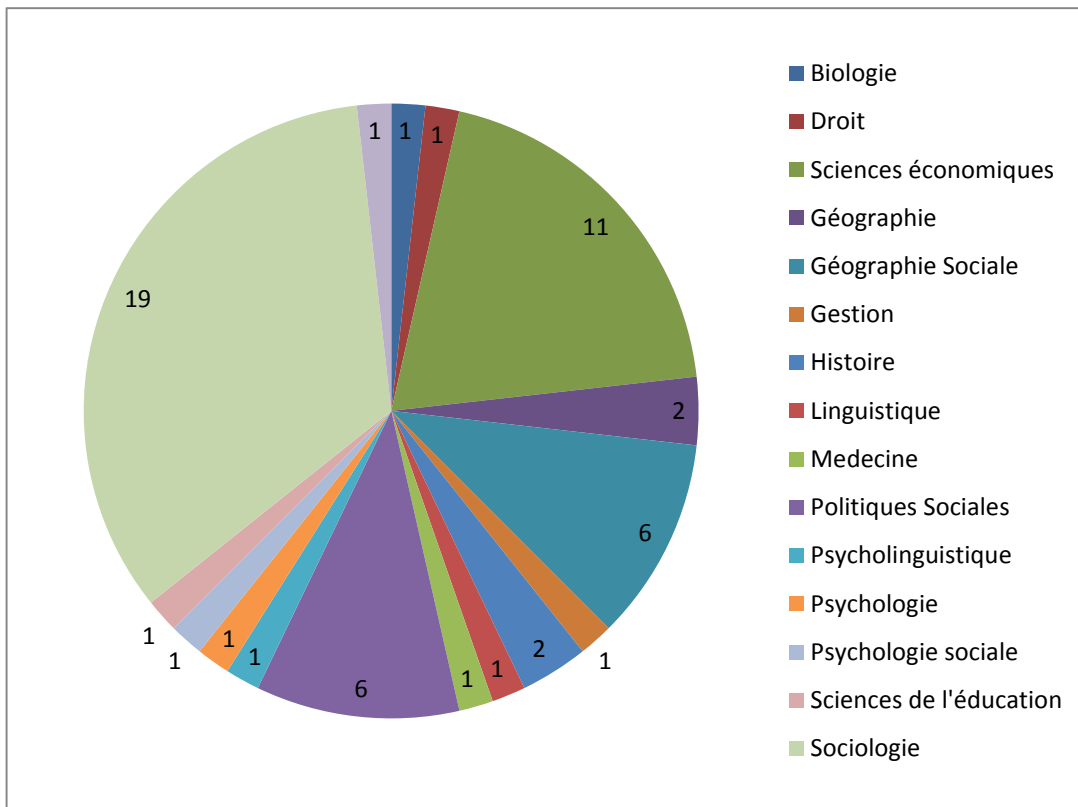
Depuis le lancement de la plateforme universitaire de données de Caen, il y a eu 56 utilisateurs et utilisatrices de ses services. Les publics sont assez variés, avec des

contacts plus fréquents avec les doctorants. Les disciplines sont aussi très diversifiées, même si près du tiers des utilisateurs proviennent de la sociologie et un deuxième tiers de géographie sociale et des sciences économiques. Le dernier tiers concerne les politiques sociales (IUP MSS), les sciences de l'éducation, la médecine, la biologie, le droit, l'histoire, la gestion, les STAPS, la psychologie, la psychologie sociale et la psycholinguistique. La Plateforme Universitaire de Données de Caen étant implantée au sein de la Maison de la Recherche en Sciences Humaines au sein de l'établissement de l'Université de Caen, il n'est pas étonnant de constater que la majorité de ses utilisateurs appartiennent à celle-ci. Néanmoins, il apparaît quelques cas qui ne sont pas directement de l'université de Caen mais qui ont bien sûr un lien indirect (un codirecteur de thèse ou de mémoire, ou un doctorant à la fois enseignant vacataire au sein de l'UCBN).

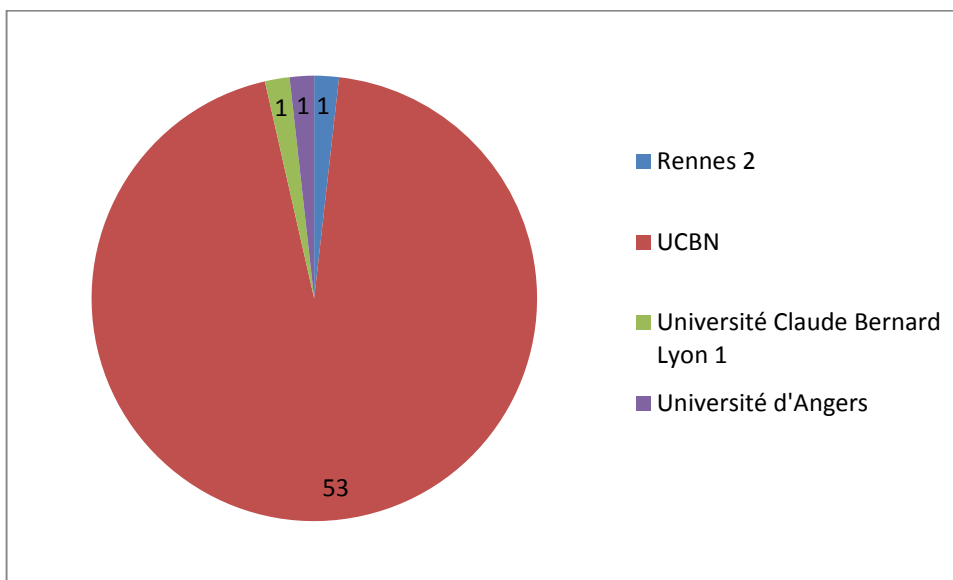
En ce qui concerne le laboratoire d'accueil pour les utilisateurs hors étudiants en master (qui ne sont pas nécessairement rattachés à un laboratoire), il n'est pas étonnant, si l'on se réfère à la répartition par discipline, de voir les laboratoires de Sociologie et de Psychologie (CERReV), de Géographie Sociale (ESO) et d'Economie et de Management (CREM) être le plus largement représentés parmi les effectifs de la plateforme.



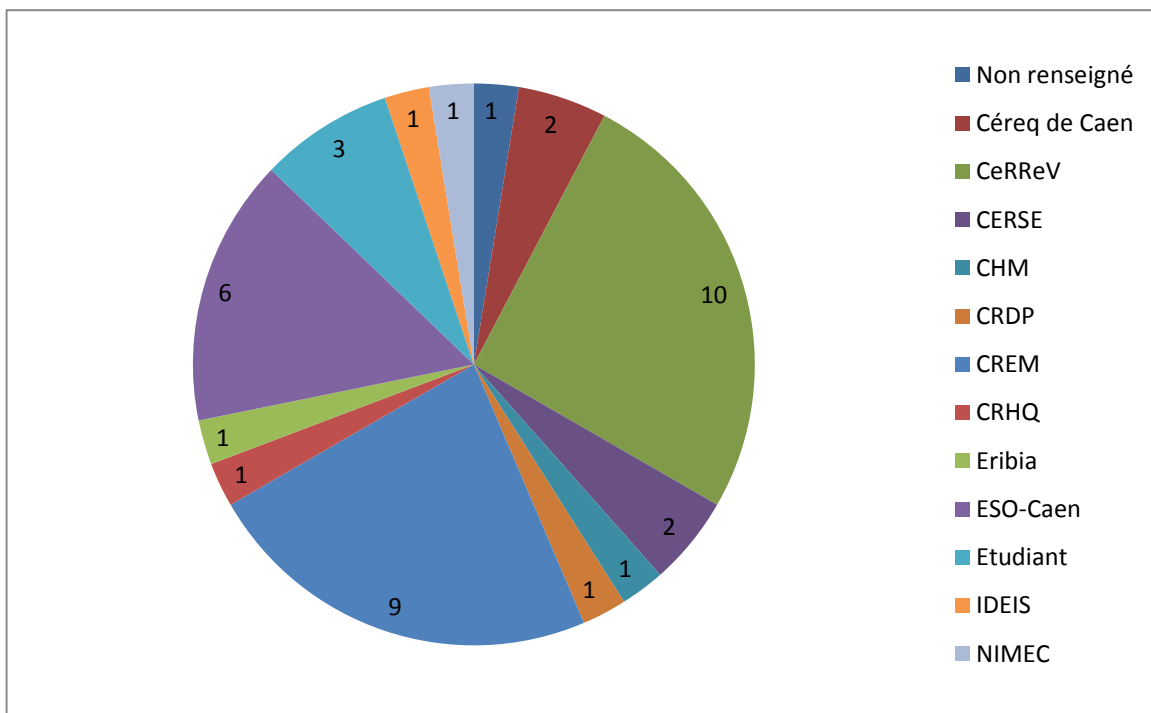
Répartition des utilisateurs de la PUDC par statut.



Répartition des utilisateurs par discipline.



Répartition des utilisateurs selon leur université de rattachement.

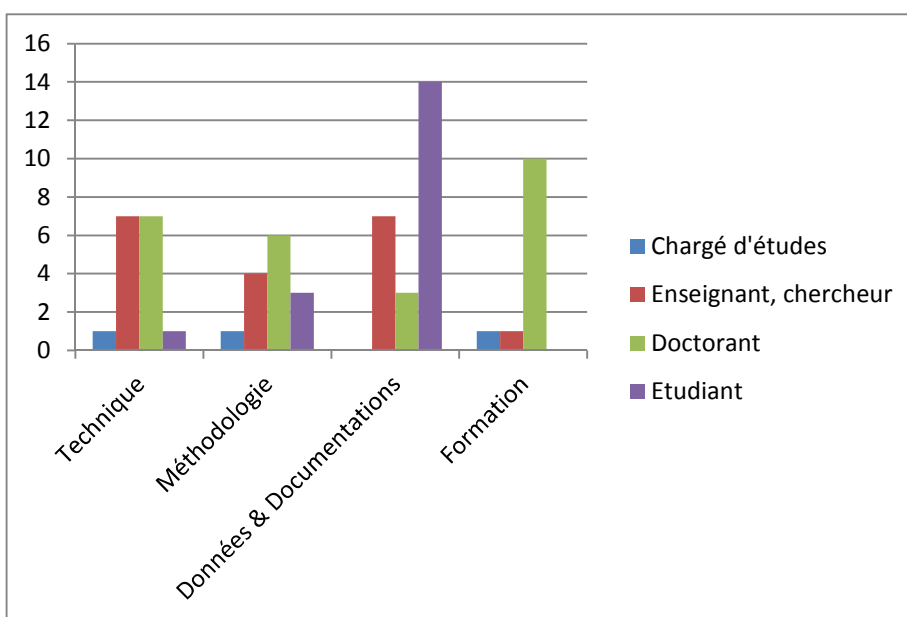


Répartition des utilisateurs de la PUDC par laboratoire d'accueil.

b. La demande

La demande exprimée concernait pour 16 d'entre eux des aspects techniques, pour 14 des considérations méthodologiques, pour 24 des données et de la documentation autour de ses données et pour 12 des formations personnalisées ou en groupe.

Si l'on s'intéresse au statut des utilisateurs, on se rend compte que les enseignants-chercheurs (y compris les chargés d'études) ont des demandes assez variées même si la formation ne semble pas être leur préoccupation principale, les doctorants sont les plus demandeurs de formation et les étudiants de données et de documentations.



Composition de la demande des utilisateurs de la PUDC par statut.

2-La formation par groupes

a. La formation SAS.

Du 19 avril 2011 jusqu'au 20 mai 2011, la plateforme a proposé à 2 doctorants en sciences économiques, une formation de 10 séances de 3 heures pour une initiation à l'utilisation du logiciel SAS.

Du 15/06/2011 jusqu'12/07/2011, 2 doctorants et 1 docteur en sciences économiques se sont vu proposer la même formation sur une dizaine de séances, mais les emplois du temps des uns et des autres étant difficilement compatibles. Nous avons dû nous contraindre à stopper cette formation au bout de 4 séances. Cependant, nous avons pu reprendre une session à partir du 12/10/2011 jusqu'au 23/11/2011 pour 11 séances de 3 heures. Elle était composée de deux doctorants en sciences économiques (dont un de la session précédente) et un doctorant en sociologie. Une autre session est ouverte depuis le début du mois de mars.

Suite à ces formations, le bureau de la formation des personnels de la direction des ressources humaines de l'Université de Caen a contacté la plateforme afin qu'elle puisse prendre en charge la formation des personnels de l'université de Caen quant à l'utilisation du logiciel SAS.

b. La formation Statistique (en projet)

Suite à une distribution massive des plaquettes de la PUDC, des doctorants en histoire sont spontanément venus pour signifier leur besoin en termes de statistiques. Ainsi, nous avons pour projet de proposer une formation orientée « outils statistiques » pour ces doctorants.

3-L'appui individualisé

La plateforme de données propose aussi un accompagnement individualisé adapté à chaque demande.

a. Les besoins en données quantitatives du réseau Quetelet et les documentations

Les demandes des utilisateurs en données quantitatives du point de vue de la plateforme

Vingt-quatre utilisateurs sont venus à la plateforme pour avoir des renseignements sur l'existence de données utiles pour leur projet de recherche. Si pour trois d'entre eux elles concernaient des données internationales, les vingt-et-une autres recherchaient des enquêtes françaises. Parmi ceux-là, quatre n'ont pas eu de réponses positives quant

à l'existence d'une enquête appropriée pour leur thématique de recherche. Pour le reste, il leur a tous été proposé au moins une enquête de la statistique publique.

Les demandes des bas-normands du point de vue du réseau Quetelet.

Il y a eu trente-cinq commandes de bases de données au réseau Quetelet des personnes localisées en Basse-Normandie du début de l'année 2011 jusqu'à maintenant. Une quinzaine d'entre elles se sont faites via la plateforme universitaire de données de Caen et sept autres se sont faites sans l'aide de la plateforme.

	Economie	Epidémiologie, Santé	Géographie	Sociologie	Total général
Via la PUDC	2		4	10	16
Hors PUDC	9	10			19
Total général	11	10	4	10	35

Tableau des commandes passées auprès du réseau Quetelet durant l'année 2011 (jusqu'au 12/03/2012). Source Réseau Quetelet.

b. L'assistance technique et méthodologique

L'une des prérogatives de la plateforme concerne l'assistance des chercheurs, enseignants, doctorants et étudiants en Master dans la mise en œuvre d'une enquête quantitative de la conception du protocole d'enquête jusqu'à l'exploitation des résultats. Pour ce faire, la PUDC propose une assistance individualisée. Ainsi 20 utilisateurs avaient des demandes d'aide individualisée. Pour 3 d'entre eux, cela concernait l'élaboration et la mise en forme d'un questionnaire. Cinq ont sollicité le soutien de la plateforme pour mettre en forme leur base de données (ou simplement convertir dans des formats de fichier plus adaptés). Treize pour une assistance logicielle quelconque et six pour l'aide à la mise en œuvre de résultats.

4- La diffusion de séminaires

Trois séances du Séminaire « Méthodes et Modélisation Statistiques en Sciences Sociales » ont été diffusées via le système de visio-conférence ceux du 29/03/2011, du 31/05/2011, du 18/10/2011 et du 31/01/2012. La première et la deuxième séances ont concerné deux personnes, la troisième, trois personnes et la dernière une seule.

5- L'archivage de données.

La PUDC a commencé un travail d'archivage d'une enquête régionale menée par le Centre associé au Céreq de Basse-Normandie.

Durant les premiers temps de la mise en place de la PUDC, nous avons entrepris de mettre en forme, pour pouvoir les diffuser par la suite, les données du « PAnel de Suivi des ELèves » produites localement par le Céreq de Caen. Ces données qui interrogent les parcours scolaires, les projets (professionnels et scolaires) et l'orientation scolaire sont constituées de 3 vagues. La PUDC a effectué la mise en forme des données et la constitution du dictionnaire des variables pour la 1^{ère} vague. Cependant, des questions d'ordre organisationnel et institutionnel concernant la diffusion de ces données restent encore en suspens. C'est pourquoi nous avons interrompu la mise en forme de la seconde vague du PASTEL et celle d'une enquête concernant les lycéen-ne-s atypiques en termes de genre dans les formations techniques et professionnelles des lycées et centres de formation d'apprentis en Basse et Haute-Normandie (enquêtes produites par Clotilde Lemarchant).

IV-Les projets pour 2012

a. Communication

Le travail déjà engagé visant à faire connaître l'existence de la plateforme de données doit être poursuivi. D'une part, des étudiants nouveaux en master2 et doctorat doivent être informés chaque année des services proposés. D'autre part, il est toujours utile à la fois de rappeler l'existence de cette plateforme aux collègues déjà informés, et vers les collègues et laboratoires qui n'ont pas encore été formellement rencontrés. Ainsi, l'accent sera mis en 2012 sur les disciplines du droit et d'histoire (CRHQ, centre de recherche d'histoire quantitative).

b. Développer le travail en réseau

Il s'agit de développer le travail en réseau avec les différentes plateformes universitaires existantes ou en cours de création et le CASD (Centre d'accès sécurisé distant), en accord avec les grandes lignes politiques du GIS « réseau Quetelet ».

c. L'observatoire de la mixité

La plateforme universitaire de diffusion de données est adossée à différents projets de recherche. Le premier est de constituer un groupe d'observation « genre, trajectoires de formations et ressources humaines à l'université de Caen Basse-Normandie », animé par M. Pierre Sineux, Vice-président de l'université (CEVU), et 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, l'Observatoire de l'université de Caen / SUOIP ; 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 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 Hémerly, master 2 psychologie sociale, encadré par Patrice Georget, MCF.

- Formation : les étudiants, l'orientation et le genre. Sur ce second domaine, le travail est en cours (rencontre des différents partenaires, définition des indicateurs). La PUDC, en lien avec l'observatoire de l'université, peut coordonner les modalités institutionnelles et techniques afin de faire un état des lieux sur les formations supérieures à l'épreuve du genre à l'université de Caen.

Le CEVU se 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 relève d'une démarche d'établissement, et sera inscrit dans le contrat d'établissement. Un document doit être remis aux conseils d'université pour juin 2012.

d. L'ANR INEDUC

La PUDC est partenaire d'une ANR coordonnée par Patrice Caro de ESO-CAR Cereq Caen sur les relations entre activités sur temps libre et parcours scolaires chez les collégien-ne-s.

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'éducation renvoie ici aux acquisitions qui résultent des interactions entre l'individu et son environnement physique et humain, englobant tant les transmissions par action volontaire que par imprégnation.

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. Il s'agit plus précisément d'étudier la dimension spatiale de ces contextes, que ce soit par le vecteur analytique des distances métriques ou par celui des distances sociales et culturelles.

Faute d'actions préventives ou réparatrices efficaces, ces différences se traduisent en inégalités qui, cumulées dans l'espace et dans le temps chez une partie des jeunes, peuvent mettre en péril la cohésion sociale. Identifier les différents mécanismes de la construction de ces inégalités et leurs liens peut fournir des outils théoriques utiles pour la promotion de l'épanouissement et du développement du potentiel de chaque

jeune afin qu'il puisse devenir à la fois acteur de son propre développement et du développement social.

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 plateforme est partenaire de l'équipe caennaise de ce projet dans les phases d'investigation documentaire et d'analyse quantitative (analyse secondaire de données de statistiques publiques et production de données propres par questionnaires).

- e. La formation statistique avec l'Ecole doctorale des sciences humaines et sociales

Dans le prochain exercice, il est prévu de proposer une formation orientée « outils statistiques » pour les doctorants qui en éprouveraient le besoin. Pour cela, il conviendra de se rapprocher de l'école doctorale (cf III.2.b).

- f. La formation SAS auprès des personnels UCBN

Pour l'année 2012, nous espérons pour concrétiser l'approche que nous avons opérée auprès du bureau de la formation des personnels de la direction des ressources humaines pour la tenue de séances de formation pour l'utilisation du logiciel SAS (II.2.a).

- g. L'archivage des données des enquêtes en préparation

Comme indiqué dans précédemment dans le chapitre III.5, nous prévoyons aussi pour l'année 2012 de continuer la préparation des données déjà en notre possession pour une possible diffusion via le réseau Quetelet.