



EDITO

Bertrand Jouve,

Directeur Adjoint Scientifique à l'Institut des Sciences Humaines et Sociales (CNRS)

Le développement du numérique a entraîné dans nos communautés scientifiques une hégémonie de la « donnée scientifique numérique » : elle s'échange facilement, rapidement et se prête souvent à des traitements automatiques. La construction d'une Europe de la recherche est devenue indissociable d'une Europe des données, plongée dans le « big data », le calcul intensif ou l' « open access ».

La conséquence est une transformation des pratiques des chercheurs dans tous les champs scientifiques et les sciences humaines et sociales ne font pas exception. Il est nécessaire d'accompagner et de faciliter cette transformation en proposant des services adaptés, transnationaux, européens, distribués et participatifs, et donc conçus selon les meilleurs standards internationaux : normes d'interopérabilité des données, traitements automatiques de grands corpus, services d'indexation des bases de données, numérisation intelligente, sauvegarde pérenne, etc. L'enjeu est important car c'est à cette condition que nous donnerons aux chercheurs relevant des sciences humaines et sociales des facilités à participer activement à la réflexion sur les grandes questions de notre temps, que ce soit autour du défi climatique, du vieillissement de la population ou des nouveaux modes

de sociabilité en réseau pour ne citer que quelques exemples.

Quels sont les verrous actuels dans le domaine des Humanités Numériques ? Quel type d'édition est compatible avec le « big data » et l' « open access » ? Quelles valeurs ont les données numériques de recherche en sciences humaines et sociales ? Comment développer le partenariat entre le milieu académique et les entreprises ? Comment articuler les initiatives européennes, nationales et régionales ? Ces questions seront abordées lors des rencontres Nouvelles collaborations scientifiques autour des infrastructures du numérique : France-Europe-Québec qui se tiendront à Toulouse les 26 et 27 novembre prochains.

L'ambition de ces rencontres est aussi d'ouvrir la voie à la participation de nouveaux grands sites scientifiques français à ces dynamiques européennes et internationales autour du numérique et de la question des données, aussi bien dans leurs dimensions académique, qu'économique et citoyenne, en mettant en perspective les enjeux de société qui y sont liés.

Data without Boundaries : quels progrès pour l'accès aux données très détaillées de la statistique publique en Europe ?

Conduit dans le cadre du 7ème programme cadre, le projet européen Data Without Boundaries (DwB) a travaillé à lever les obstacles qui rendent difficiles pour les chercheurs la mobilisation des données très détaillées ou sensibles des grandes bases de la statistique publique, au-delà des frontières nationales.

Retour sur ce projet mené de 2011 à 2015 : nouveaux outils disponibles, progrès accomplis et propositions faites.

Les grandes bases de la statistique publique constituent des ressources extrêmement riches pour la recherche et la société. Les dispositifs d'accès sécurisés aux données indirectement identifiantes ou sensibles, tels que le CASD pour la France, permettent désormais plus facilement de les utiliser. Mais les obstacles persistent pour les projets qui mobilisent des sources de plusieurs pays : barrières juridiques, diversité des procédures d'autorisation, dispersion de l'information, et dispositifs sécurisés peu adaptés au travail collaboratif entre chercheurs, sont autant de difficultés rencontrées par les chercheurs désireux d'accéder à des données confidentielles au-delà des frontières nationales.

Bâtir une coopération au niveau européen pour faciliter les accès transnationaux

S'inspirant des modèles fructueux de coopération entre acteurs de la recherche et producteurs de statistique publique qui ont permis d'ouvrir l'accès à ces données au plan national (Réseau Quetelet, INSEE et services statistiques ministériels pour la France, UKDA et ONS pour la Grande-Bretagne), DwB a rassemblé 28 partenaires, banques de données pour la recherche, centres de recherche méthodologique, et instituts nationaux de statistique publique de 14 pays, sous la coordination scientifique de Roxane Silberman, directrice du Réseau Quetelet. Organisés par domaines de compétence, les groupes de travail de DwB ont travaillé à harmoniser les usages et pratiques et à bâtir une coopération au niveau européen entre le CESSDA, le réseau européen des banques de données dont PROGEDO est partenaire, et le système statistique européen coordonné par Eurostat.

Vers un réseau européen de centres sécurisés

Traçant les grandes lignes d'une feuille de route européenne vers une plus grande harmonisation,

des recommandations ont été formulées sur les standards pour les métadonnées, les critères d'accréditation des chercheurs pour l'accès aux données confidentielles, et les principes du contrôle des sorties pour les données traitées via des accès sécurisés.

L'architecture proposée pour un réseau européen de centres sécurisés (European Remote Access Network, EuRAN) a pour objectif de faciliter le travail collaboratif d'équipes de chercheurs sur des sources hébergées dans des centres sécurisés différents, tout en garantissant la nécessaire protection de la confidentialité. Le niveau de sécurité très élevé permet à un tel réseau de fonctionner dans les cadres juridiques existant. Un premier contrat entre l'IAB en Allemagne (service statistique du ministère de l'emploi) et le UKDA en Grande-Bretagne a permis de vérifier l'efficacité de la notion de cercle de confiance entre partenaires européens. Cet accord, fondé sur l'équivalence des dispositions juridiques et des procédures de sécurité, permet aux chercheurs accrédités en Grande-Bretagne d'utiliser les données très détaillées d'IAB sans avoir à se rendre en Allemagne. Dans la lignée de cette expérimentation, des accords équivalents sont en négociation pour

la France et les Pays-Bas, amorces de coopération européenne entre centres sécurisés.

Un projet pilote réalisé par le CASD a permis de démontrer concrètement la faisabilité informatique d'un tel réseau au delà de liaisons bilatérales. (cf encadré).

De nouveaux outils pour l'exploration des micro-données de la statistique publique en Europe

Très concrètement également, les partenaires du projet ont travaillé sur les solutions pour améliorer l'information et la documentation sur les données nationales. Deux nouveaux outils sont nés de cette réflexion :

- **CIMES** (Centralising and Integrating Metadata from European Statistics), un catalogue de métadonnées de base, en anglais et au format international DDI, sur les données disponibles pour chacun des pays ainsi sur que les procédures d'autorisation et les liens avec les diffuseurs. Pour cet outil, créé et développé par les partenaires du Réseau Quetelet - ADISP en coopération avec le CASD, la participation française a été centrale.
- **MISSY** (Microdata Information System) un programme qui organise la documentation jusqu'au niveau fin des variables pour les enquêtes européennes Eurostat et facilite leur usage pour la recherche comparative. Le GESIS, la banque de données allemande, a adapté son outil à la structure de ces enquêtes pour la mise en œuvre de cette solution très appréciée par Eurostat.

Réseau sécurisé européen, une réalité pour demain ?



Mis en place par le CASD, un Proof of Concept de réseau sécurisé européen a été développé dans le cadre de DwB à partir de serveurs installés sur 3 nœuds européens (CASD/France, GESIS/Allemagne, et DESTATIS/Allemagne) et un nœud central. Présenté en mars 2015, lors de la seconde édition du European Data Access Forum à Luxembourg, ce projet a permis de démontrer concrètement la faisabilité d'un tel réseau.

Une coopération renforcée

La coopération entre les partenaires, créée par le cadre du projet, est sortie renforcée du travail mené en commun. Cette dynamique s'est étendue aux communautés parties prenantes des questions d'accès aux données, mobilisant dans les discussions les banques de données pour la recherche, les instituts nationaux de statistique européens et impliquant également directement les chercheurs - accès gratuits aux centres sécurisés, sessions de formation aux enquêtes Eurostat - et les centres intéressés par la création d'accès sécurisés à distance qui ont pu visiter les installations existantes.

Des coopérations nouvelles ont vu le jour entre banques de données et instituts de statistique, là où elles n'existaient pas ou peu. D'ores et déjà, un accord est en cours de discussion pour formaliser une coopération entre le CESSDA et Eurostat, et pérenniser le European Data Access Forum, plate-forme européenne permettant de confronter régulièrement les besoins des différents acteurs et les avancées réalisées. Des projets sont en cours pour la construction d'un réseau sécurisé européen dans le cadre du prochain appel H2020.

En savoir plus

DwB:
<http://www.dwbproject.org/>

CIMES:
<https://cimes.casd.eu/>

MISSY:
<http://www.gesis.org/missy/>

DONNEES

Nouveautés 3ème trimestre 2015

CESSDA Fr / Réseau Quetelet - La vie après un cancer : des données disponibles

Les données de l'enquête sur la vie deux ans après un cancer (VICAN2) sont désormais mises à disposition de la communauté scientifique, via l'Adisp. A cette occasion, une convention a été signée entre l'EHESS (pour le compte de PROGEDO/Réseau Quetelet), l'Institut national du cancer et l'Inserm.

Cette enquête réalisée en 2012, porte sur 4 349 personnes dont le cancer a été diagnostiqué deux ans avant l'étude. Elle dresse un panorama de la vie après un cancer, sur le plan médical, psychologique, social et professionnel et poursuit deux objectifs principaux :

- l'évaluation du second plan cancer, dont les actions prévues comprenaient la réalisation de cette enquête,
- la détermination de la nature, la prévalence et la temporalité des facteurs physiques, psychologiques et sociaux susceptibles d'altérer les conditions de vie, la réintégration sociale et la (ré)insertion professionnelle des personnes atteintes d'un cancer.

Les données de VICAN2 sont accessibles via l'Adisp.

ESS - Appel pour les modules de la 9ème édition

L'appel à projet pour les modules de la 9ème édition de l'Enquête Sociale Européenne est en ligne. Les équipes de recherche multi-nationales en sciences sociales ayant un attrait pour les démarches comparatives sont invitées à soumettre des projets de modules de questions qu'elles souhaiteraient voir intégrer à la 9ème édition de l'ESS.

Les dossiers de candidature, à rendre pour le 11 janvier 2016, devront mettre en avant la pertinence des propositions pour des recherches transnationales, le calendrier de travail, et les qualifications des membres de l'équipe soumissionnaire concernant le design de questionnaire.

L'appel à proposition complet est disponible sur le site de l'ERIC ESS : www.europeansocialsurvey.org/

GGP - Des syntaxes pour le traitement des données

Pour soutenir les recherches faites sur les données de GGP, l'équipe de coordination propose de mutualiser les routines de traitement de ces données, via son site web : la première syntaxe proposée en ligne permet de convertir toutes les dates de GGS dans le format année/mois, plus facile d'utilisation.

Désireuse d'étendre ce catalogue avec l'aide des utilisateurs, l'équipe de GGP invite tous les chercheurs à contribuer à ce projet de partage de syntaxes; les programmes disponibles sur le site seront crédités du nom de l'auteur.

Si vous souhaitez renforcer la recherche menée par la communauté, envoyez vos syntaxes en format STATA ou SPSS à ggp@nidi.nl



EVENEMENTS

Nouvelles collaborations scientifiques autour des infrastructures du numérique : France-Europe-Québec

Organisées à **Toulouse les 26 et 27 novembre 2015**, les rencontres Nouvelles collaborations scientifiques autour des infrastructures du numérique : France-Europe-Québec donneront une vision synthétique des dynamiques en cours concernant le développement des infrastructures de recherche en sciences humaines et sociales.

Plus d'informations sur le site www.univ-tlse2.fr/msh



Vos données sécurisées au cœur de la datascience

Première conférence autour des données organisée par le CASD, Vos données sécurisées au cœur de la datascience interrogera les nouvelles potentialités offertes par la sécurisation accrue de l'accès aux données.

Rendez-vous à **Paris le 8 décembre 2015**, en présence de Thomas Piketty, Antoine Frachot, Roxane Silberman et Philippe Donnay.

Nombre de place limité, pré-inscriptions et informations sur : casd.eu/fr/event2015

GGP - conférence utilisateurs

La 3ème conférence des utilisateurs de GGP aura lieu les 30 novembre et 1er décembre prochains au Vienna Institute of Demography. Cette rencontre internationale biennale permettra aux chercheurs travaillant avec les données et/ou bases de GGP, de présenter leur travail, et d'en discuter les approches méthodologiques et les résultats empiriques.

ACTU DES DEPARTEMENTS

CESSDA-FR - Nouveau style pour le site de l'Adisp

Pour la rentrée, le site de l'Adisp fait peau neuve et s'adapte désormais aux appareils mobiles. Plus moderne et convivial, il reprend la description de l'ensemble des enquêtes et bases de données diffusées, sans oublier les Produits Sur Mesure (PSM) ou les projections de population réalisées à partir de l'outil Omphale. L'ensemble des informations concernant l'Adisp, ses partenaires, et l'accès aux données sont également accessibles depuis son nouveau menu.

ESS - Bien être social et personnel des Européens

L'Enquête Sociale Européenne a publié fin juin un livret sur les principaux résultats de la 6ème édition, consacré au bien être social et personnel des Européens. Une version française a été éditée et mise en ligne début septembre. Ce livret reprend les principaux résultats issus du module rotatif relatif au bien-être social de la 6ème édition de l'ESS, qui déconstruit le concept complexe et global de « bien-être » et en fournit une analyse unique et comparative.

Ce document est accessible sur le site français de l'enquête : <http://ess.sciencespo.com>

SHARE - une rentrée active

La prochaine journée annuelle de SHARE se tiendra à Dauphine le **25 Novembre 2015 de 9h00 à 18h00**. Cette journée, centrée chaque année sur une thématique spécifique, a pour objectif de porter à la connaissance d'un public plus large que la seule communauté scientifique, des résultats récents et de favoriser les rencontres entre chercheurs, institutionnels et acteurs de terrain. Des informations sur cet événement seront très prochainement disponibles sur le site de SHARE France.

Toutes les informations disponibles sont sur le site share.dauphine.fr

PUD - Bientôt Dijon

L'automne est florissant pour les plates-formes universitaires de données : après Nantes, la MSH de Dijon organise la demi-journée de lancement de la nouvelle PUD de **Dijon lundi 9 novembre de 14h à 16h30**.

Données disponibles, accès sécurisés, et offre des PUD seront présentés lors de cette rencontre.

Plus d'informations à venir sur le site <http://msh-dijon.u-bourgogne.fr>

PUD - Nouvelle venue à Nantes

A l'occasion du lancement de sa plate-forme universitaire de données, la MSH Ange Guépin de Nantes a organisé vendredi 18 septembre 2015 une demi-journée de rencontres avec la communauté de recherche locale. Une quarantaine de personnes a assisté aux présentations qui ont permis de cadrer l'environnement de cette nouvelle plate-forme : TGIR Progedo, données de l'Adisp, activité des PUD. Une matinée riche en échanges et prometteuse pour l'activité de cette PUD.