

Les enjeux de la mise à disposition de données d'enquêtes qualitatives à caractère personnel. Retour sur l'expérience beQuali

Guillaume Garcia, Emilie Groshens

▶ **To cite this version:**

Guillaume Garcia, Emilie Groshens. Les enjeux de la mise à disposition de données d'enquêtes qualitatives à caractère personnel. Retour sur l'expérience beQuali. Le cadre juridique applicable aux traitements de données à caractère personnel, journée d'études (org. M. Moritz), Apr 2017, Lille, France. hal-02874203

HAL Id: hal-02874203

https://sciencespo.hal.science/hal-02874203

Submitted on 18 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Les enjeux de la mise à disposition de données d'enquêtes qualitatives à caractère personnel.

Retour sur l'expérience beQuali

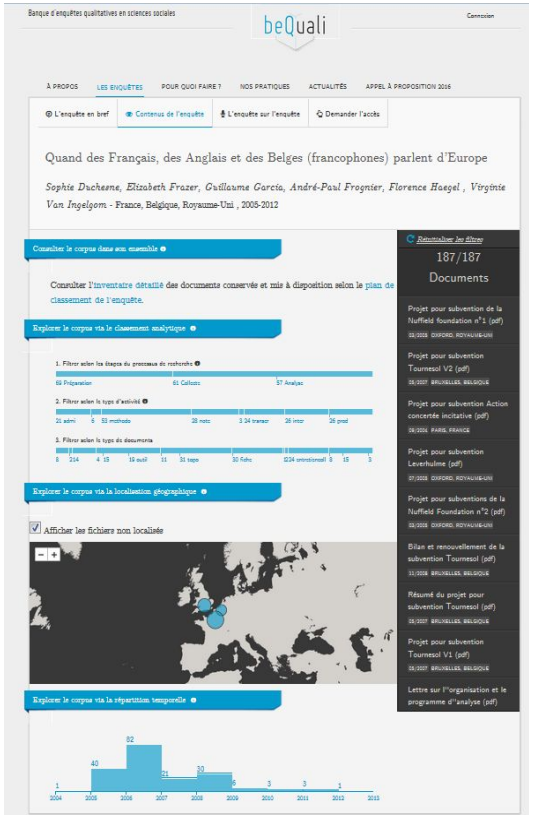
Journée d'étude « Le cadre juridique applicable aux traitements de données à caractère personnel »
Lille, vendredi 28 avril 2017

Guillaume Garcia, Emilie Groshens - Centre de données socio-politiques
https://cdsp.sciences-po.fr

beQuali, une banque d'enquêtes qualitatives

SciencesPo
CENTRE DE DONNÉES SOCIO-POLITIQUES

Un **équipement national** développé au CDSP



Un **site web** de mise à disposition d'enquêtes en SHS



Plusieurs **enjeux :**

- la préservation
- la visibilité des travaux
- **la réutilisation pédagogique et scientifique**

Diffuser des données qualitatives de sciences sociales : de l'idée à la réalité

Des données complexes à fort enjeu éthique

- Paroles (récits de vie, opinions) et propriétés sociales
- Méthodes de recueil : l'immersion, la confiance (observation participante, entretien en face à face)
- Enquête = Corpus hétérogène de données
  - collectées et co-construites (enregistrements, transcriptions, notes de terrain)
  - variées (administratives, brutes/traitées, d'analyse)
  - multi-supports (carnet papier, audio, vidéo)



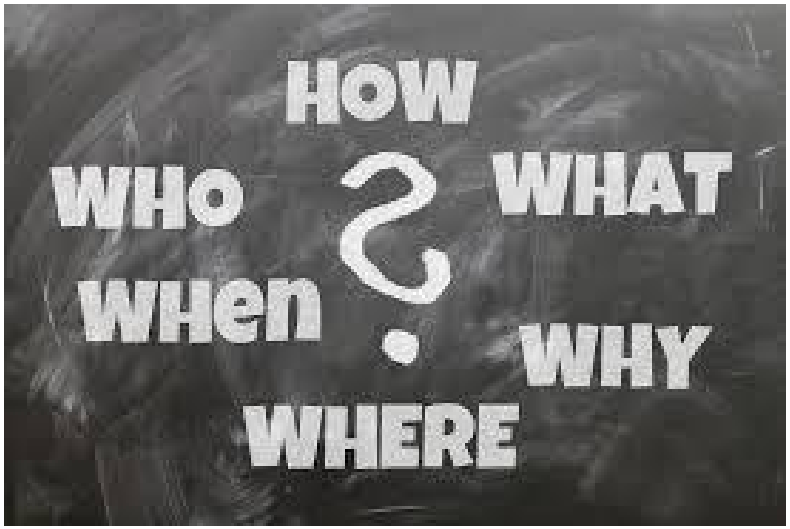
Des tensions scientifiques et juridiques fortes

- **D'une part**, respecter le Droit, protéger les enquêtés (Code patrimoine, PI, CNIL, respect vie privée, droit image*)
- **D'autre part,** garantir le potentiel de réutilisation (exhaustivité, finesse, contextualisation des données)

Bricoler en ‘terra incognita’

Des pratiques peu répandues en France

- **Faible culture** du partage et de la réutilisation des données qualitatives
- **Craintes et réserves** des chercheurs : rupture d'anonymat ou de confidentialité
- **Peu d’écrits** et de retours d’expériences sur l’anonymisation





Un flou juridique : “comment faire ?”

- pas de précisions pratiques dans la loi
- pas de recettes dans les préconisations institutionnelles (CNIL)
- absence de recommandations de type déontologique

Dans un contexte scientifique où la donnée a une valeur importante
(résultat d’une recherche, validité/preuve, coût de production, reproductibilité difficile)

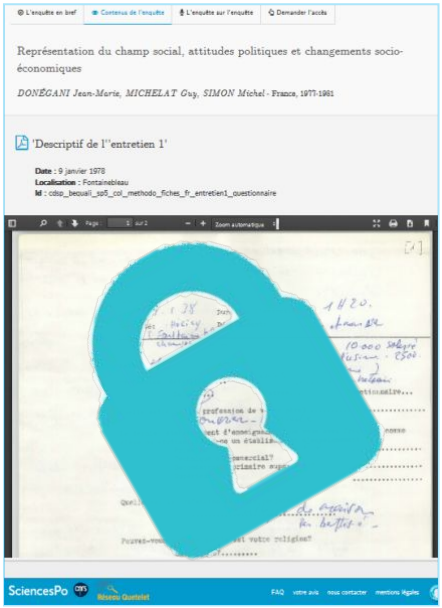
beQuali

Inventer au cas par cas une gestion du risque “adaptée”

Nos solutions à l’échelle globale

- Contrat pour le déposant
- Accès restreint à la communauté scientifique
- Contrat d’engagement (éthique) à chaque demande
- Anonymisation ou demandes de consentement
- Collaboration avec le CIL CNRS
- Respect de l’intégrité de l’oeuvre

Réseau Quetelet



Verrouillage du site

1 protocole d’anonymisation /enquête

- Collaboration du chercheur
- Règles générales : équilibre risque / besoin scientifique
- Suppression des identifiants directs (nom, adresse, tel)
- Repérage et traitement des identifiants indirects (croisement d’info, anecdotes, métadonnées)
- Numérique : hyperonymes

Exemple de l'enquête d'Agnès van Zanten "Choisir son école"

- **Objet :** comprendre les motivations du contournement de la carte scolaire
- **Terrain ethnographique :** 1999-2005, 4 villes de région parisienne étudiées 2 à 2 par communes limitrophes
- 400 documents : 150 transcriptions, une dizaine de notes terrain
- 2 populations différentes : parents d'élèves & acteurs de l'École
- Globalement : ***critère géographique > critère sociographique***

Types d'enquêtés	Données conservées	Données supprimées	Justification du choix au regard du risque
Parents d'élèves	Ville, quartier, école, Profession, Délégué classe, Membre du CA	ID (identifiants directs), *Nationalité*, *Employeur (nom, lieu),* Nom des relations, *Nom de l'association*	Réputation (contournement de normes, opinions politiques, religieuses..). Fonctions éphémères. Difficulté d'accéder aux archives des écoles.
Représentants d'associations de parents d'élèves	Ville, école, Association / Responsabilité	ID, *Nom de l'association (si "président")*	Risque d'identification si Président.
Acteurs scolaires (profs, directeurs, inspecteurs) ; Élus locaux ; Agences immobilières	Ville, *Poste occupé*	*ID (sauf inspecteurs et maires),* Ecole	Impossibilité d'anonymiser les responsables hauts placés. Risque mineur : parole contrôlée. Choix de privilégier le critère géographique.

Exemple de l'enquête d'A. Gotman "Dilapidation et prodigalité"

- **Objet :** dilapidation d'héritages, prodigalité et surendettement (jusqu'aux mesures de sauvegarde de justice)
- **Terrain :** 1990-1993, en région parisienne et en province
- 300 documents : 60 entretiens, 30 questionnaires, 30 notes de dossiers de tutelle, quelques notes d'observation
- **Groupes étudiés :**
  - Dilapidateurs
  - Informateurs sur la dilapidation
  - Familles surendettées
  - Membres des "débiteurs anonymes"
  - "Experts" (assistantes sociales, psychiatres, juristes et juges des tutelles)
  - Réutilisation d'anciennes enquêtes ("petits dilapidateurs" & accédants à la propriété)
- Protocole d'anonymisation modulé selon les composantes de l'enquête
- Anonymisation affinée des témoignages de dilapidateurs (sensibilité)
- Globalement : ***critère sociographique > critère géographique***

Les problèmes pratiques

- Une enquête = un ensemble documentaire & un maillage informationnel complexe

- Principes d'anonymisation
  - Respect de l'intégrité des archives (réversibilité / copies)
  - Pas de transformation ni de falsification de l'information : suppression (post-it) ou hyperonyme

- Etendue des opérations d'anonymisation
  - Matériaux (transcriptions d'entretien, notes d'observation)
  - Documents de préparation & d'analyse
  - Métadonnées (diversité des métadonnées)
  - Protection des enquêtés, personnes citées & chercheurs

- Outils de contextualisation pour compenser le retrait d'information
  - Tableaux de correspondance enquêtés / pseudos
  - Documentation du protocole dans l'"enquête sur l'enquête"





En conclusion : compromis, gestion du risque et limites

- L'anonymisation : **un compromis entre des opérations, des compétences et des contraintes** multiples
  - Scientifiques
  - Juridiques
  - Archivistiques
  - Techniques

- L'anonymisation : des arrangements **au cas par cas**
  - Dépend de la participation du chercheur
  - Impossible d'élaborer un protocole générique
  - Bonne gestion du risque

- **Les gardes fous** de l'anonymisation : ne pas travestir les données

- **Des chantiers à mener**
  - Enquêtes ethnographiques
  - Données audio
  - Données vidéo

Merci de votre attention. Des questions ?

guillaume.garcia@sciencespo.fr

emilie.groshens@sciencespo.fr

contact@bequali.fr



www.bequali.fr

Photos : Pixabay, Wikimedia