
Des métadonnées et paradonnées pour mieux exploiter des données d'enquêtes pour la recherche

Valentin Brunel¹, Paul Colin¹, Alina Danciu¹, Quentin Gallis¹, Emilie Groshens², Lucie Marie¹, Geneviève Michaud¹, Baptiste Rouxel¹, Tom Villette¹

¹ Centre de données socio-politiques (CDSP)

² Maison méditerranéenne des sciences de l'homme (MMSH)

hal-03432005, v1

Une équipe pluridisciplinaire

Une vingtaine d'ingénieurs experts en :

- Production d'enquêtes
- Traitement de données
- Documentation, diffusion de données
- Développement d'applications

Le CDSP valorise
les données
de la recherche
pour la communauté
académique



Catalogue de données du CDSP

- Enquêtes en sciences humaines et sociales traitées et contextualisées par les ingénieurs du CDSP depuis 2005, dans le respect des principes FAIR et des normes internationales en vigueur (DDI, TEI...)
- Des enquêtes produites à Sciences Po ou dans d'autres institutions françaises et internationales, issues de méthodes quantitatives ou qualitatives
- Principaux thèmes : série des enquêtes pré- et post-électorales françaises (1958, 1962, 1978, 1988, 1995, 1997, 2002, 2007, 2012), science, environnement, santé, sexualité, pratiques culturelles, relations entre les générations, etc.

En guise d'introduction...

Principaux standards et outils pour la documentation des données et paradonnées

- Data Documentation Initiative (DDI)
- [Dataverse](#)
- Nesstar
- Colectica
- Outils et plateformes *in-house*

Métadonnées



Source: <https://bit.ly/35lPIIK>

Beuster, Benjamin, & Fry, Jane. (2020, November 30). What can DDI do for you? An introduction to the DDI. 12th Annual European DDI User Conference (EDDI20). Zenodo. <https://doi.org/10.5281/zenodo.4298302>



Source: <https://bit.ly/2lvHGIt>

Beuster, Benjamin, & Fry, Jane. (2020, November 30). What can DDI do for you? An introduction to the DDI. 12th Annual European DDI User Conference (EDDI20). Zenodo. <https://doi.org/10.5281/zenodo.4298302>

Plan

- (Bonnes) Pratiques et outils de documentation
- La confidentialité des données au défi du partage de données en auto-dépôt
- Retour expériences utilisateurs sur le catalogue du CDSP
- Paradoxaux

Pratiques et outils de documentation

<https://www.menti.com/trc2ygexk4>



Une petite introduction

Documenter une enquête...

Pourquoi faire ?

Comment faire ?



Une petite introduction

Un objectif principal : la réutilisation du jeu de données !

Des difficultés :

Une situation d'enquête non vécue

Un terrain non expérimenté

Une interprétation erronée des données hors contexte

Des données inintelligibles, illisibles, incomplètes...



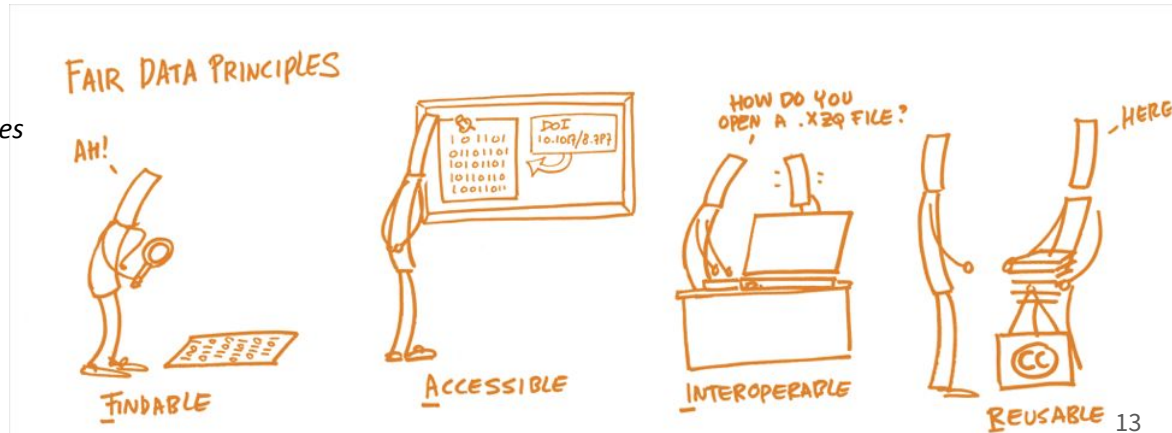
Une petite introduction

Comment garantir le potentiel de réutilisation ?

Du FAIR (Findable, Accessible, Interoperable, Reusable) et du FLAIR (quels usages potentiels ?)

- Reusable* :
- ▢ Licence d'utilisation
 - ▢ Sources
 - ▢ Métadonnées riches, pertinentes
 - ▢ Standards (discipline)
- = **contextualisation**

Et des outils...



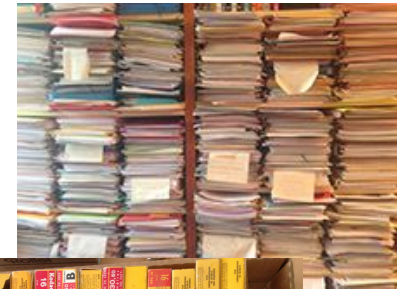
Qu'est-ce qu'une enquête ?



Des données collectées et co-construites à partir d'un terrain (une population enquêtée) sur une période délimitée pour répondre à une question de recherche.

Quali / Quanti :

- des corpus denses (préparation, collecte, analyse)
- des supports et formats multiples
- des données personnelles ou sensibles
- des archives collectives, personnelles, institutionnelles



Le périmètre de l'enquête

Comment sélectionner les documents du corpus ?

Quelles sont les bornes spatio-temporelles de l'enquête ?
(définir l'unité "enquête")

Quelles sont les ressources pertinentes pour une réutilisation ?
(versions, brouillons, méthode, codes...)

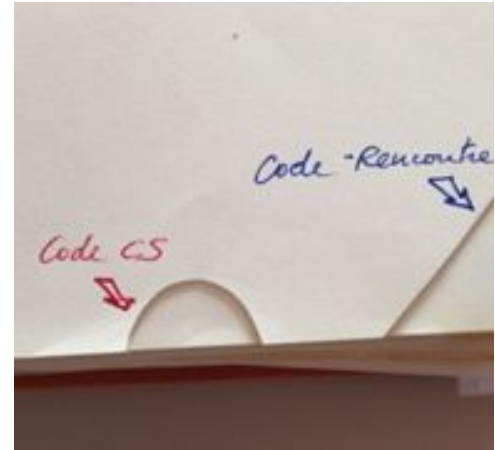
Quels sont les droits applicables aux données ?
(consentements, non communicables / sensibles, édition, propriété intellectuelle...)



Le périmètre de l'enquête

Trois éléments essentiels :

- Contexte
- Données
- Documentation



Le périmètre de l'enquête

Le contexte

= tout ce qui éclaire sur les objectifs et conditions de production
(Pourquoi ? Qui ? Pour qui ? Avec qui ? Avec quel financement ? Quand ? Où ?)

Exemples :

Documents de cadrage : projet de recherche, compte-rendus, carnet de bord

Contexte scientifique : bibliographie, enquêtes préparatoires

Méthode :

Il est aussi essentiel d'échanger, si l'on n'est pas producteur, avec la personne ayant réalisé l'enquête



Le périmètre de l'enquête

Les données

= toutes les informations prélevées sur le terrain étudié
(opinion, image, son, observation, document)

Exemples :

Questionnaire, entretien, photo, vidéo, carnet de terrain, documentation collectée

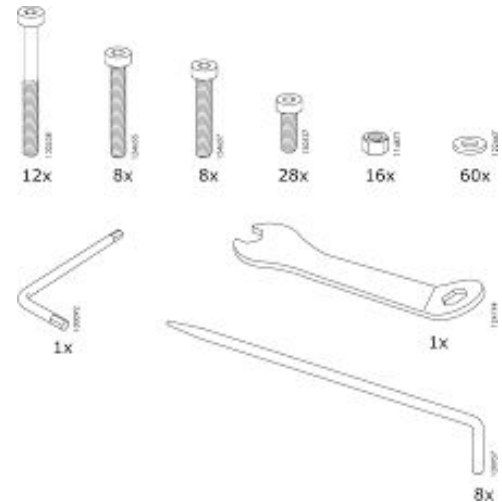
Différentes versions plus ou moins brutes ou traitées

Des transferts de supports (son → texte, questionnaires → variables)

Méthode :

Sélectionner les données utiles à la réutilisation avec l'aide du producteur.

Disposer des données complètes et les rendre compréhensibles.



Le périmètre de l'enquête

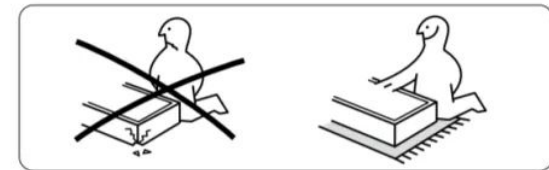
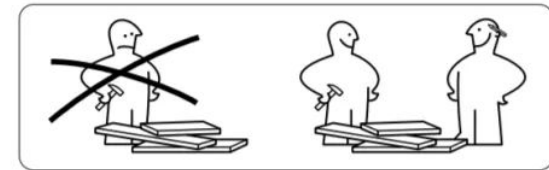
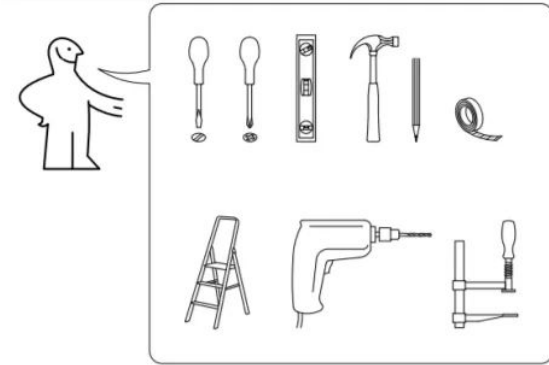
La documentation

= l'essentiel pour pouvoir exploiter les données de terrain
(Comment précisément ? Les méthodes de collecte et d'analyse)

Exemples :

Quali : guides d'entretien, tableaux de suivi, courriers de sollicitation, annotations, notes, synthèses, tables de pseudos, consignes de transcription, etc.

Quanti : questionnaire, codebook, filtres éventuels, notes sur les recodages, etc.



Quelques critères de sélection

Un exemple d'enquête, le volet quantitatif du *Panel de Caen*



Quelques critères de sélection

ives > enquetes_en_cours > cdsp_bq_s14_Bidart > Corpus > Original > Docs_disquedur

Recherche

Action Outils Paramètres

Nom	Taille	Type de fichier	Date de modification
Aentretiens		Dossier	01/09/2021 10:54:28
Aguide		Dossier	04/03/2021 09:48:56
Bentretiens		Dossier	18/02/2021 11:14:32
Bguide		Dossier	01/10/2020 14:27:17
Calendriers		Dossier	13/10/2020 14:02:54
Centretiens		Dossier	18/03/2021 15:00:33
Cguide		Dossier	04/03/2021 10:06:29
codage		Dossier	01/09/2021 11:13:19
Dentretiens		Dossier	18/02/2021 11:21:55
Dguide		Dossier	06/10/2020 09:28:42
Eentretiens		Dossier	04/03/2021 14:49:06
Eguide		Dossier	18/02/2021 11:50:26
Fenquête		Dossier	04/03/2021 09:24:43
Fentretiens		Dossier	14/11/2019 10:37:47
Fguide		Dossier	04/03/2021 09:32:58

15 éléments

panel_de_caen > fichiers_base

Rechercher dans : fichiers_base

Nom	Modifié le
A-Contacts.xls	26/10/2012 06:35
A-liensTot.xls	14/01/2017 07:41
A-QuestionnaireEgo.XLS	15/09/2016 00:27
A-Relations.xls	15/11/2010 05:03
A-Relations_geo.xls	08/11/2019 11:35
B_Villes_Alters.xls	10/11/2012 04:01
B-Contacts.XLS	26/10/2012 13:46
B-Emploi_2.xls	01/02/2011 05:07
B-liensTot.xls	26/10/2012 09:09
B-QuestionnaireEgo.xls	28/04/2008 02:24
B-Relations.xls	28/08/2007 03:39
B-stages.xls	17/10/2002 03:44
C_Villes_Alter_tot.xls	03/08/2013 23:48
C-Contacts.xls	27/10/2012 05:37
C-Emploi_2.xls	02/02/2011 06:20
Code-liensFinal.doc	22/11/2006 07:26
Code-questionnaire.doc	28/10/2005 02:04
C-QuestionnaireEgo.xls	18/07/2017 03:03
C-Relations.xls	28/08/2007 03:55
C-RelationsTot.xls	03/08/2013 23:45
C-stages.xls	08/09/2010 07:36
D-City Alter3.xls	29/07/2013 03:41
D-Contacts.xls	29/07/2013 03:42
D-Emplois.xls	08/09/2010 07:39
depot_bidard_101020.zip	06/10/2020 13:52
D-liensTot.xls	09/11/2012 05:41
D-QuestionnaireEgo.xls	20/04/2012 03:05
D-Relations-mai07.xls	24/05/2007 03:42
D-Stages.xls	08/09/2010 07:40

Quelques critères de sélection

enquetes_en_cours > cdsp_bq_s14_Bidart > Corpus > Original > Docs_disquedur > codage

Action Outils Paramètres Recherche

Nom	Taille	Type de fichier	Date de modification
archives		Dossier	14/02/2020 10:28:1
codagev1		Dossier	04/03/2021 16:19:5
codagev2		Dossier	14/11/2019 10:45:5
codagev3		Dossier	27/10/2020 10:02:2
codagev4		Dossier	14/11/2019 10:45:5
codagev5		Dossier	14/11/2019 10:45:5
codagev6		Dossier	15/06/2021 15:20:3
ABC-Code-Cerdes.doc	44 KB	DOC Fichier	07/11/2002 11:31:5
ABC-code-liens.doc	120 KB	DOC Fichier	11/07/2003 17:58:0
ABC-Code-Réseau.doc	24.5 KB	DOC Fichier	30/04/2003 12:34:4
ABC-Code-Stock.doc	21 KB	DOC Fichier	06/03/2007 17:16:4
ABCD_Coordonnees_geogr_Egos_121215.xls	1.2 MB	XLS Fichier	15/12/2012 08:50:5
ABCD_Coordonnees_geogr_Egos_130812.xls	1.2 MB	XLS Fichier	17/09/2013 19:53:1
ABCD_Densite_Siena.xls	833.5 KB	XLS Fichier	22/07/2010 06:47:3
ABCD_EgoAlterDyads_Couples_130422.sav	2.1 MB	SAV Fichier	22/04/2013 15:53:4
ABCD_networks_Siena.xls	833.5 KB	XLS Fichier	22/07/2010 06:45:1

70 élément(s)

File Station

enquetes_en_cours > cdsp_bq_s14_Bidart > Corpus > Original > Docs_disquedur > codage > codagev1

Action Outils Paramètres Recherche

Nom	Taille	Type de fichier	Date de modification
A_AgeEnf.xls	402.5 KB	XLS Fichier	03/11/2010 14:11:36
A_Cerdes.xls	185 KB	XLS Fichier	03/11/2010 15:22:06
A_LiensInfluence3PourSPSS.xls	112 KB	XLS Fichier	29/08/2007 10:28:10
A_ReseauDensiteX.csv	137 KB	CSV Fichier	17/06/2008 15:11:30
A_ReseauDensiteX.xls	484 KB	XLS Fichier	04/03/2021 16:19:55
A_Villes_Alters.xls	388 KB	XLS Fichier	05/08/2013 08:38:02
A-City Alter-Alter.xls	2.9 MB	XLS Fichier	24/11/2010 16:31:18
A-City Alter.xls	303 KB	XLS Fichier	08/11/2010 11:54:30
A-City Alter3.xls	261 KB	XLS Fichier	05/08/2013 18:25:42
A-Code-EGO.doc	83 KB	DOC Fichier	15/05/2003 11:38:12
A-Contacts.xls	1 MB	XLS Fichier	26/10/2012 15:35:16
A-liens.xls	1.4 MB	XLS Fichier	17/07/2003 16:41:48
A-liensTot.xls	2.1 MB	XLS Fichier	14/01/2017 16:41:34
A-Question.XLS	125 KB	XLS Fichier	26/01/2007 12:55:10
A-QuestionnaireEgo.XLS	157 KB	XLS Fichier	15/09/2016 09:27:02
A-Relations_geo.xls	227 KB	XLS Fichier	08/11/2019 20:35:47

20 élément(s)

Quelques critères de sélection

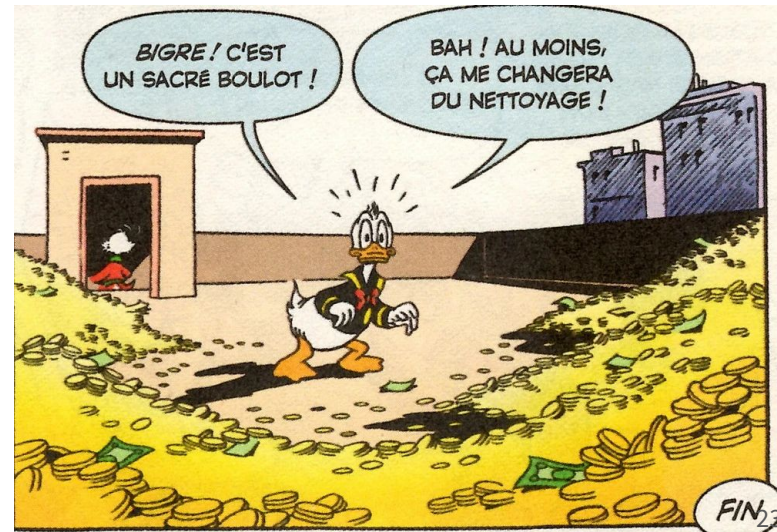
Les données sont en état de traitement par la chercheuse...

Corpus extrêmement exhaustif, parfois redondant, où les différentes étapes de la recherche se sont sédimentées.

Parole de la chercheuse :

“si on n’est pas dans ma tête, on n’y comprend rien”.

Comment s’y repérer ?





	Nom	Type	Largeur	Décimales	Libellé	Valeurs	Manquant	Colonnes	Align	Mesure	Rôle
1	wave_ego	Chaîne	3	0		Aucune	Aucune	4	Gauche	Nominal	Entrée
2	EGO	Numérique	1	0		Aucune	Aucune	5	Droite	Echelle	Entrée
3	ego_number	Chaîne	2	0		Aucune	Aucune	8	Gauche	Nominal	Entrée
4	name_ego	Chaîne	10	0		Aucune	Aucune	8	Gauche	Nominal	Entrée
5	wave	Chaîne	1	0		Aucune	Aucune	6	Gauche	Nominal	Entrée
6	wave_num	Numérique	8	0		Aucune	Aucune	5	Droite	Echelle	Entrée
7	vague_annee	Numérique	8	0		Aucune	Aucune	5	Droite	Echelle	Entrée
8	filiere	Numérique	1	0	{1, bac ES}...	Aucune	8	Droite	Nominal	Entrée	
9	vague	Chaîne	1	0		Aucune	Aucune	5	Gauche	Nominal	Entrée
10	XXX_EGO	Numérique	8	2		Aucune	Aucune	8	Droite	Echelle	Entrée
11	gender_ego	Numérique	1	0	{1, garçon}...	Aucune	8	Droite	Nominal	Entrée	
12	filiere_scolaire_ego	Numérique	1	0	{1, bac ES}...	Aucune	11	Droite	Nominal	Entrée	
13	social_origin_ego	Numérique	1	0	{1, sup}...	Aucune	8	Droite	Nominal	Entrée	
14	birth_date_ego	Numérique	2	0		Aucune	Aucune	12	Droite	Nominal	Entrée
15	Age_Ego_Ci4	Numérique	8	0	{1, 17 to 21 ...	Aucune	13	Droite	Echelle	Entrée	
16	Age_ego_cl3	Numérique	8	0	{1, 17 à 20 ...	Aucune	9	Droite	Echelle	Entrée	
17	Age_Ego_Ci2	Numérique	8	0	{1, 17 to 32 ...	Aucune	13	Droite	Echelle	Entrée	
18	Age_Ego	Numérique	8	0		Aucune	Aucune	6	Droite	Echelle	Entrée
19	activity_ego_cl5	Numérique	8	0	{1, en étude...	Aucune	8	Droite	Echelle	Entrée	
20	activity_ego_cl3	Numérique	8	0	{1, études}...	Aucune	14	Droite	Echelle	Entrée	
21	activity_Ego_3cl	Numérique	8	0	{1, studies}...	Aucune	16	Droite	Nominal	Entrée	
22	SITEGO1	Numérique	1	0	{0, autre}...	Aucune	8	Droite	Nominal	Entrée	
23	SITEGO2	Numérique	1	0	{0, autre}...	Aucune	8	Droite	Nominal	Entrée	
24	work_or_not	Numérique	8	1	{0, pas en ...	Aucune	13	Droite	Echelle	Entrée	
25	diplom_ego_cl7	Numérique	1	0	{0, no diplo...	Aucune	13	Droite	Nominal	Entrée	
26	level_educ_ego_cl7	Numérique	1	0	{1, college}...	Aucune	8	Droite	Nominal	Entrée	
27	level_educ_ego_cl4	Numérique	8	2	{1,00, inf à ...	Aucune	10	Droite	Echelle	Entrée	
28	PCS_ego_INSEE	Numérique	2	0		Aucune	Aucune	12	Droite	Echelle	Entrée
29	PCS_ego_student_first	Numérique	8	0		Aucune	Aucune	14	Droite	Echelle	Entrée
30	PCS_ego_cl10	Numérique	8	0	{10, agricult...	Aucune	12	Droite	Echelle	Entrée	
31	PCS_ego_cl6	Numérique	8	0	{20, agricul...	Aucune	14	Droite	Echelle	Entrée	

Vue de données

Vue des variables

Quelques critères de sélection

Les variables sont de différents types.

- variables “de base” : l’information essentielle et sans laquelle on ne peut effectuer aucun traitement.
 - identifiant
 - informations socio-démo
 - variables d’intérêt (dans le cas du panel de Caen, réseau)

Quelques critères de sélection

Les variables sont de différents types.

- variables “recodées” : l’information a été traitée par la chercheuse, afin de permettre l’analyse
 - exemple des PCS : au moins trois variables
 - variables sur la structure du réseau (typologies)
 - recodages de questions à choix multiple

Quelques critères de sélection

Les variables sont de différents types.

- variables “de spécialiste” : la recherche est encore en cours, ou très pointue
 - les pourcentages en classe de certaines caractéristiques de réseau (arêtes, isolés, etc.)
 - médiane de distance au sein du réseau
 - variables de localisation à l’intérieur des réseaux

Quelques critères de sélection

Conserver les données de base.

Sélectionner, en accord avec le producteur, les variables recodées les plus pertinentes. C'est du cas par cas !
=> exemple : pour les PCS, garder la variable la plus complète, à partir de laquelle il est possible de recréer d'autres classes si besoin.

=> pour la typologie de réseaux, garder la typologie effectuée, mais pas les typologies secondaires qui en sont dérivées

En règle générale, supprimer les variables trop techniques ou trop circonstanciées (variables dichotomisées, recodages de circonstance, recherche en cours...)

Quelques critères de sélection... en quali

Est-ce que le document

- est diffusable (droit et déontologie) ?
- est unique ?
- est compréhensible ?
- renseigne sur la démarche ?

Exemple pour l'enquête *Le temps de travail domestique*

(D. Fougeyrollas et D. Chabaud-Rychter, 1985 - corpus numérisé)

- choix de diffuser les transcriptions d'entretien en ODT (texte) + PDF (image)
- les originaux comportant des annotations, corrections et les codes couleur de la méthode d'analyse

ménage

Courses

bricolage - jardinage - rapp à la technique

Cuisine

Couture - raccommodage - ouvrages

lavage - repassage

travail

maï

enfants

parents ~~ma~~ amis - ~~ma~~ voisins

budget

école - ψ

médecine - hygiène - ~~coiffeur~~ beauté...

déplacements

horaires

organ. de l'espace domestique

coiffeur - beauté - entretien de soi-même

Femme de ménage : ----

Bottes =

R : Le problème c'est un peu le samedi. Bon ben je suis obligée de l'emmener là bas à l'école, et puis je perd beaucoup de temps parceque... le temps d'aller le conduire et puis faut que je vienne prévenir mes parents qu'il est à l'école parceque il suffit que j'aie un ennui le midi, soit que j'ai été retardée ou n'importe, ils ignorent si le gamin il est encore en classe, enfin en principe il y est mais on sait jamais ce qui peut se passer. Alors si jamais le midi ils ne me voient pas revenir il faut qu'ils aillent rechercher le petit à l'école. Alors je reviens ici à neuf heures trente à peu près, quelques fois faut que j'aille à l'école faire les courses. Bon ben le temps de revenir, j'ai juste le temps de faire manger et il est onze heures et demie, faut que j'aille le rechercher.

Q : Hum.

R : Bon ben si y'a des jours où je travaille pas, c'est pareil il faut... Je vais le conduire le matin, à cinq heures, il faut que j'aille le rechercher.

Q : Et même le samedi, vous déjeunez chez vos parents?

R : Non, non. Le samedi non. Ben le samedi, le petit part qu'à neuf heures enfin, on part d'ici à neuf heures moins vingt. L'école c'est qu'à neuf heures. Alors il déjeune avec nous.

Q : Et vous le... (?)

R : Ben le samedi on dîne tous les trois ensemble, parcequ'on est là. Enfin, un samedi sur deux, c'est pareil, mon mari travaille mais il rentre à une heure et demie alors, comme je reviens de l'école il est midi et demie à peu près, on attends que lui pour manger... Je donne à manger au petit quand même, parcequ'il est habitué à manger le midi, s'il doit attendre de manger jusqu'à une heure et demie heu... Bon ben il mange Je lui donne à manger tout seul, puis heu, ça lui permet de venir remanger avec nous après, un enfant... Enfin c'est pour lui éviter d'aller rechercher des gateaux et des biscuits parcequ'après... il veut plus manger.

Q : Hum.

R : Je crois que... Les problèmes bon ben/ non ce qui est (?) c'est surtout pour l'enfant qu'on rencontre des obstacles... *mais on c'est fait cette vie,*

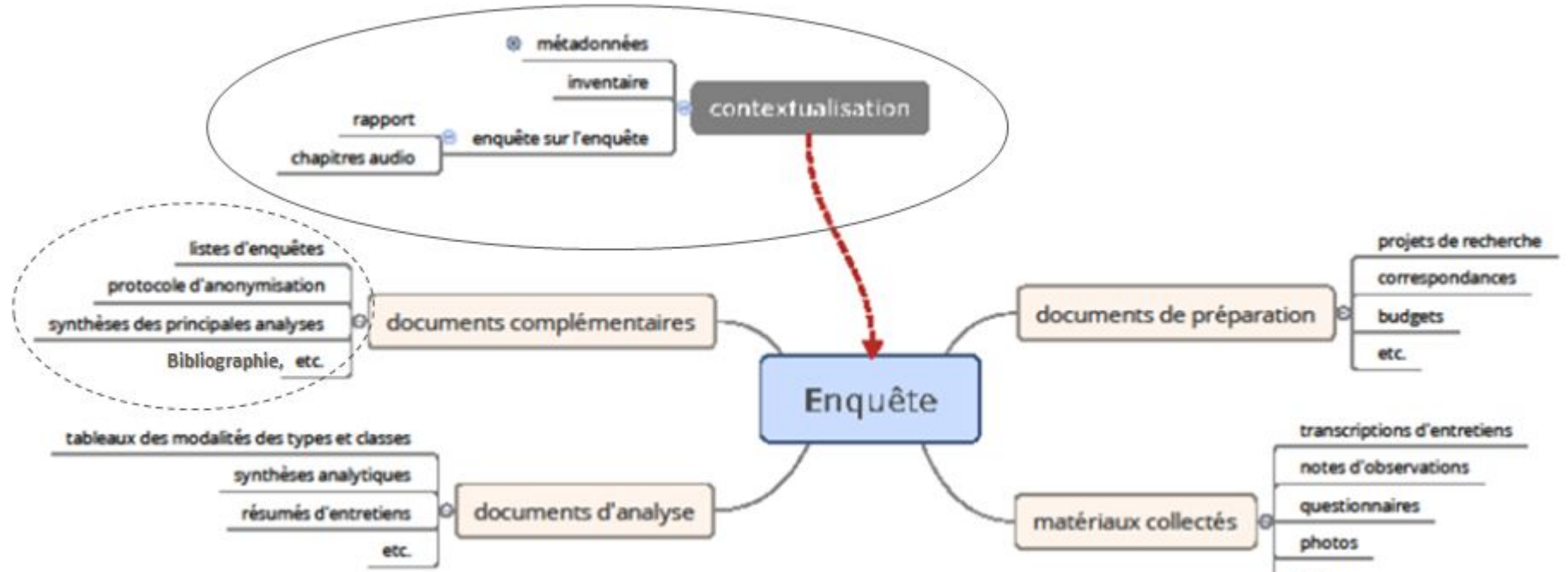
Q : Hum.

R : Parceque d'abord heu... on est obligé de se lever très tôt. Même quand

Quoi conserver ?

- **Données de terrain** produites (transcriptions d'entretien, notes d'observation, photos, enregistrements audio/video)
- **Documentation collectée** sur le terrain ou au fil de la recherche
- **Documents administratifs** (projet de recherche, courriers, missions, formulaires consentement)
- **Travail collectif** (tableau de suivi et répartition des entretiens, CR réunion, cahier collectif, consignes de transcription et d'anonymisation)
- **Documents méthodologiques** (guide d'entretien, grilles ou schémas d'analyse)
- **Traces de traitement** (surlignage, annotations, mémos et codes logiciels ou artisanaux)
- **Rédactions intermédiaires** (journal de bord, brouillons, notes, synthèses)
- **Enquêtés** (table de correspondance pseudos/entretiens, propriétés sociographiques)
- **Bibliographie** (littérature consultée et produite)

Documenter un jeu de données qualitatives



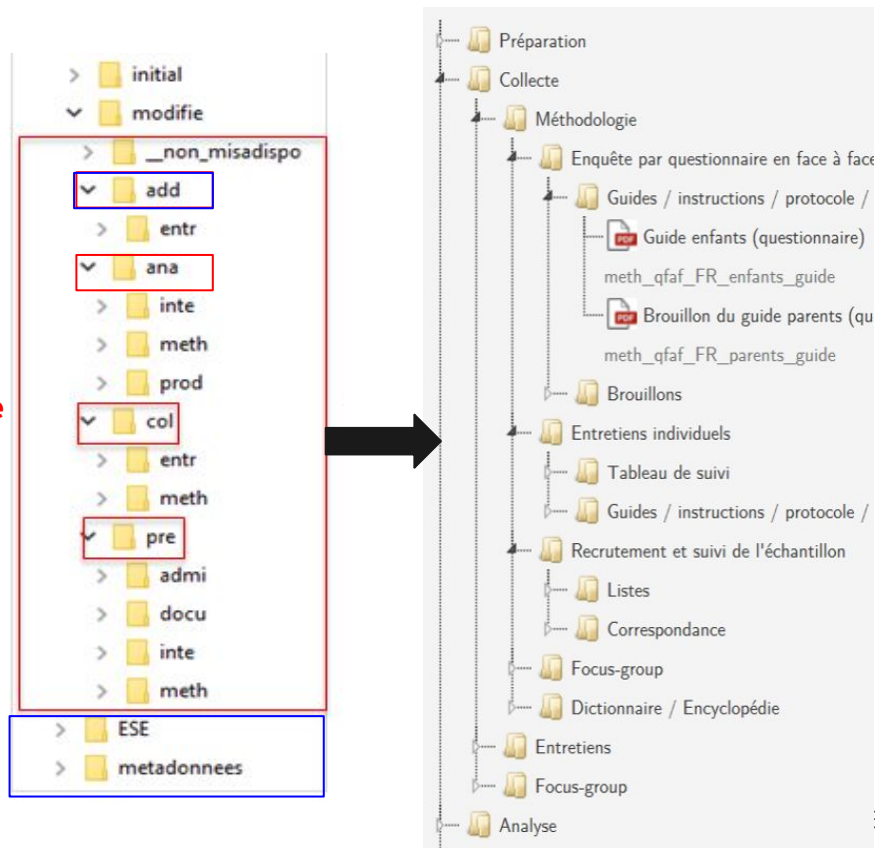
Le plan de classement : l'exemple beQuali

Organisation arborescente :

- corpus
- documentation
 - inventaire,
 - plan de classement,
 - enquête sur l'enquête,
 - documents additionnels,
 - readme, etc.

corpus de l'enquête

documentation



1e niveau Sous-fonds	2e niveau Dossier	3e niveau Sous-dossier	4e niveau Article	
[ADD] Compléments	[ADM] Documents administratifs	[ANXE] Observations annexes	[ACTES] Actes de colloques	
[ANA] Analyse	[ARCHIVES] Archives	[BRCH] Brochures	[ARTCL] Articles	
[COL] Collecte	[DOCU] Documentation rassemblée en amont ou pendant terrain	[BUDG] Budget	[AUDIO] Enregistrements audio	
[PRE] Préparation	[ENTR] Entretiens	[CODE] Codage	[BIBL] Bibliographie	
↓ Phase enquête	[INTE] Documents intermédiaires	[COMM] Communications	[BRLLN] Brouillons	
	[METH] Méthodologie	[DICO] Dictionnaire / Encyclopédie	[CARTE] Cartes	
	[OBSE] Observations	[ECHT] Recrutement et suivi de l'échantillon	[CHAPT] Chapitres	
	[PROD] Production scientifique	[ELAQ] Élaboration du questionnaire	[CMPTA] Pièces comptables	
	[QNT] Volet quantitatif	[ENRG] Enregistrements / documents multimédia	[CMPTR] Compte-rendu	
	↓ Type activité	[ENSG] Enseignement	[COURS] Cours	[COUVT] Couverture
		[EXQL] Exploitation des données qualitatives	[CPDCE] Correspondance	[EXAME] Sujet d'examen
		[EXQT] Exploitation des données quantitatives	[FOCU] Focus-group	[EXPOS] Textes d'exposés
		[FIGU] Documents figurés (images, graphique, histogrammes, tableaux etc)	[FORM] Recrutement et formation des enquêteurs	[FICHE] Fiches
		[FOCU] Focus-group	[INDV] Entretiens individuels	[GRAPH] Graphiques
[FORM] Recrutement et formation des enquêteurs		[LGRI] Littérature grise	[GUIDE] Guides / instructions / protocole / argumentaire / plan d'exploitation	
[INDV] Entretiens individuels		[POST] Entretiens post-quant	[INDEX] Index	
[LGRI] Littérature grise		[PRDQ] Presse / périodique	[LISTE] Listes	
[POST] Entretiens post-quant		[PREL] Entretiens préliminaires	[MEMOI] Mémoires / thèses	
[PRDQ] Presse / périodique		[PROG] Programmation	[NOMCL] Nomenclature	
[PREL] Entretiens préliminaires	[PUBL] Publications	[NOTES] Notes		
[PROG] Programmation	[QFAF] Enquête par questionnaire en face à face	[OFFIC] Texte juridique / réglementaire / publications officielles		
[PUBL] Publications	[QOUV] Relevés de réponses aux questions ouvertes	[OUVRG] Ouvrage		
[QFAF] Enquête par questionnaire en face à face	[TERR] Observation du terrain principal (journal, relevés)	[PHOTO] Photographies		
[QOUV] Relevés de réponses aux questions ouvertes		[PLANS] Plans / sommaires		
[TERR] Observation du terrain principal (journal, relevés)		[PRGRM] Programme de colloque / conférences		
		[QUEST] Formulaires / questionnaires		
		[RESUM] Résumés		
		[RPPRT] Rapports / synthèses		
		[TABLE] Tableau de suivi		
		[TRACT] Prospectus		
		[TRANS] Transcriptions d'entretiens		
		[VIDEO] Enregistrements vidéo		

1. PREPARATION

1.1 Documents administratifs

- 1.1.1. Financements (dossier, devis, appel à projets, formulaire, bordereau, programme de recherches...)
- 1.1.2. Budget (convention, clauses, bons de commandes, répartitions budgétaires, notifications, factures)
- 1.1.3. Mission (ordre de mission, dossier)
- 1.1.4. Correspondance (lettre, invitation, commande de livres)
- 1.1.5. Réunion (draft, pv)

1.2 Documentation

- 1.2.1. Presse (article, dossiers)
- 1.2.2. Recrutement (calendrier d'événements)
- 1.2.3. Outils (nomenclature)
- 1.2.4. Rapports non scientifiques.

1.3 Méthodologie

- 1.3.1. Recrutement (synthèse, affiche, tract, lettre, questionnaire, tableau)
- 1.3.2. Outils (scenario, encodage, codebook, grille)
- 1.3.3. Contexte matériel (cartes, logo politique, objets présentés aux enquêtés)
- 1.3.4. Réunion (organisation, logistique)
- 1.3.5. Topo (débrief, dessin)

1.4 Production scientifique

- 1.4.1. Publications (article, chapitre)
- 1.4.2. Communications (conférence, colloque, séminaire, congrès)

2. COLLECTE

2.1. Méthodologie

- 2.1.1. Outils (protocole, tableaux)

... et les noms de fichiers

Nommage court, signifiant, unique

sans espace ni caractères spéciaux
abréviations, sujet, date AAAAMMJJ, version si évolutif...

Exemples :

projet_(typedoc/sujet_)numéro ;

pre_meth_prog_cmptr_diagnostic_FR1

<https://doranum.fr/stockage-archivage/comment-nommer-fichiers/>

Les standards de métadonnées

Niveau	DDI	EAD	METS	DC
Study				
Document				
Structure				

↓ ↓ ↓ ↓

méthode scientifique corpus d'archives corpus de fichiers **fichiers**

Décrire les fichiers : le standard Dublin Core

nom de fichier (identifiant unique)

dc:identifiant

titre

dc:title

description / légende*

dc:description

date de création

dc:date

lieu

dc:coverage

format de fichier

dc:format ; dc:type

auteur ou enquêteur

dc:creator ; dc:contributor

mots clés

dc:subject

langue

dc:language

droits

dc:rights

* pour les images : IPTC, XMP

→ **des métadonnées embarquées dans les fichiers**

→ **un tableur (ODS, CSV) : inventaire**

Des référentiels et vocabulaires

Les vocabulaires contrôlés (mots clés)

- ELSST : <https://thesauri.CESSDA.eu/elsst/en/>
- DDI Alliance : <https://ddialliance.org/controlled-vocabularies>
- IDref (géographiques, Rameau, FMeSH) : <https://www.idref.fr/>
- Rameau : <https://rameau.bnf.fr/utilisation/liste>

Les référentiels producteurs

- IDref : <https://www.idref.fr/>
- ORCID : <https://orcid.org/>
- ISNI : <https://isni.org/page/search-database/>
- ROR : <https://ror.org/>

Décrire la méthode : le standard DDI

***Data Documentation Initiative* : un niveau commun aux enquêtes quanti & quali**

Language XML pour les données statistiques et d'enquêtes en sciences sociales

- Description au niveau général : <study>
 - Univers/Terrain
 - Méthode de collecte
 - Corpus
 - Modalités d'Analyse
- Description au niveau plus fin : questions & variables

Guide des balises DDI traduit : http://www.reseau-quetelet.cnrs.fr/spip/IMG/pdf/DDI_versionFR.pdf

List of recommended elements DDI 2.2 : <http://www.ddialliance.org/sites/default/files/ddi-lite.html>

Exemple de notice DDI sur beQuali

Général	
DDI Id	cdsp_bequali_sp5
Nom	Représentation du champ social, attitudes politiques socio-économiques
Résumé	Enquête réalisée entre 1978 et 1980 sous la direction de Donégani, Guy Michelat et Michel Simon pour le Comité d'organisation des recherches appliquées sur le développement économique et social (CORDES) et faisant écho à la recherche de démarche quantitative menée en 1966 et réalisée en 1978, juste avant les élections législatives d'entretiens non directifs conduits par Claude Nédélec Donégani. L'analyse typologique qui en a découlé a fait la parution d'un rapport en 1980 intitulé 'Représentation du champ social, attitudes politiques et changements socio-économiques'.
Discipline	Science politique
Mots clés	Opinions politiques, Attitudes politiques, Religion, 1980
Pays	France
Période couverte	1978
Intervenants	
Auteurs	DONÉGANI Jean-Marie, MICHELAT Guy, SIMON Michel
Affiliations	CEVIPOF, Institut de sociologie de Lille 1
Distributeur	CDSP
Agence de financement	Comité d'Organisation des Recherches Appliquées Développement Economique et Social (CORDES)

Univers	
Niveau d'observation	National
Couverture géographique	France
Unités spatialisées	Ile-de-France, Bretagne, PACA, Auvergne
Unités d'observation	Individus
Groupes de population ciblés	Citoyens ordinaires
Dates de l'enquête	1977-1981
Méthode	
Dates de récolte des données	9 janvier 1978 - 17 mars 1978
Périodicité	Unique
Accès aux observations	La sélection des personnes interviewées a été faite de manière à constituer un échantillon représentatif des milieux sociaux.
Protocole de collecte	La même question a été posée à tous les enquêtés : "nous parlions de tout ce que représente pour vous aller faire aux prochaines élections?"
Technique de collecte	Entretiens non directifs
Corpus	
Nombre de documents	266

Corpus	
Nombre de documents	266
Types de documents	Documents textuels ; transcriptions fines paraverbal
Nombre d'observations	63 entretiens
Analyse	
Transcription	Transcriptions fines avec indication du paraverbal
Anonymisation	Effectuée par les producteurs eux-mêmes
Analyse	Analyse typologique
Langue des données	Français
Langue de la documentation	Français
Publications associées	Donegani Jean-Marie, Michelat Guy, Simon Michel, « <i>Déterminants des organisations symboliques et comportementales dans la recherche française de sociologie</i> , 1985 Volume 26, n° 1, pp. 37-41. Donegani Jean-Marie. <i>La liberté de choisir son lieu de résidence</i> , 1980

Exemple de notice DDI sur Dataverse

[data.sciencespo](#) > [Banque de données du CDSP](#) > [données qualitatives](#) > [Les conditions de vie professionnelle des femmes enceintes \[1985-1988\]](#)

[✉ Contact](#) [🔗 Partager](#)

[🌐 Publier](#)

[🔗 Lien](#)

[✎ Modifier](#)



Les conditions de vie professionnelle des femmes enceintes [1985-1988]

Brouillon

Devreux, Anne-Marie, 2021, "Les conditions de vie professionnelle des femmes enceintes [1985-1988]", https://doi.org/10.21410/dshs_2016/y568-nx23, data.sciencespo, DRAFT VERSION [?](#)

[☰ Citer le jeu de données](#)

Pour en savoir plus, consultez le document [Data Citation Standards](#).

Dataset Metrics [?](#)

0 téléchargement [?](#)

Description [?](#)

Enquête réalisée par Anne-Marie Devreux au Centre de sociologie urbaine entre 1985 et 1988. Elle s'intéresse à l'articulation entre la vie familiale et la vie professionnelle dans le cadre d'une problématique production/reproduction.

professionnelle et le moment où elles en sont éloignées. Cette étude a également révélé les réactions du milieu professionnel face à une grossesse devenant progressivement visible et rappelant l'appartenance de la salariée à une vie extérieure à l'entreprise. L'analyse du fonctionnement des rapports sociaux de sexe et de la division sexuelle du travail constitue le cadre d'analyse.

Sujet ?

Social Sciences

Mot-clé ?

SOCIOLOGIE, TRAVAIL, CONGÉ DE MATERNITÉ, RAPPORTS DE TRAVAIL, ALLOCATIONS PARENTALES, CARRIÈRE PROFESSIONNELLE, EMPLOI DES FEMMES, CHOIX PROFESSIONNEL, RAPPORTS DE TRAVAIL, ÉQUILIBRE TRAVAIL-VIE PRIVÉE, VIE FAMILIALE, MILIEU FAMILIAL, FAMILLES, CONGÉ PARENTAL, DROIT AU TRAVAIL

Related Publication ?

Devreux, Anne-Marie, "Maternité et travail : le prix de l'autonomie pour les femmes", In L'autonomie sociale aujourd'hui, actes du colloque de Biviers, organisé par le Centre d'Etudes des Pratiques Sociales (CEPS) et le Centre de Recherche sur l'Epistémologie et l'Autonomie (CREA) en novembre 1983, Grenoble, Presses universitaires de Grenoble, 1985, pp.477-486.

Remarques ?

Il existe également des communications non publiées et des articles de vulgarisation scientifique dans différents magazines. La liste complète des publications est disponible dans le jeu de données


Fichiers

Métadonnées

Conditions d'utilisation

Versions

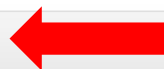
 Ajouter ou modifier des métadonnées

 Exporter les métadonnées ▾

Citation Metadata ▾

Geospatial Metadata ▾

Social Science and Humanities Metadata ▲



Unité d'analyse 

Individual

Univers 

Individus, femmes actives enceinte. Majoritairement des femmes actives enceintes appartenant à quatre groupes socioprofessionnels (cadres et professions intellectuelles supérieures, professions intermédiaires, employées, ouvrières). Le mari est parfois présent lors des entretiens et intervient. Il y a également quelques entretiens avec des assistantes sociales, des sages-femmes et des chefs de service.

Méthode temporelle 


Other

Méthode d'échantillonnage 

Nonprobability.Purposive

Taille de l'échantillon cible 


71

Mode de collecte 


Interview.FaceToFace; Recording; Transcription; CompilationSynthesis; Summary

Type d'instrument de recherche 

InterviewSchemeAndThemes

Caractéristiques de la collecte de données 

Entretiens semi-directifs. Entretiens uniques sauf pour 25 enquêtées qui ont été rencontrées une seconde fois. La plupart des entretiens a été réalisée au domicile des enquêtées par Anne-Marie Devreux et également par Irène Jonas, enquêtrice professionnelle, spécialiste des enquêtes de sociologie. Les enquêtées ont été sélectionnées par les Caisses d'allocations familiales (Paris, et l'une dans l'ouest de la région parisienne) sur les fichiers de bénéficiaires de l'allocation jeune enfant. L'échantillon a été constitué en respectant le critère de la catégorie socioprofessionnelle. L'échantillon des ouvrières a été complété par des contacts avec des sages-femmes des consultations de PMI ou d'hôpitaux, des assistantes sociales d'entreprises, etc. Il existe deux guides d'entretiens : l'un pour la première rencontre avec les enquêtées, l'autre lorsqu'un second entretien avait lieu. Lorsque la grossesse était très avancée lors de la rencontre, les deux guides d'entretiens étaient fondus en un seul. 71 transcriptions d'entretiens ; 26 notes d'entretiens ; 30 fiches d'observations ; 54 fiches et 2 notes d'analyse. Transcriptions fines avec indication du paraverbal. Analyse de contenu, réalisation de tableaux synthétiques croisés.

Remarques 

Nombre de documents : 301 fichiers mis à disposition (dont 71 transcriptions et 26 notes d'entretiens). Langage

Edition	<producer>	1.1.3.1	_____	<version>	2.1.6.1	Accès aux observations	<sampProc>	2.3.1.4
_____	<copyright>	1.1.3.2	Citation	<biblCit>	2.1.7	Echantillonnage	<sampProc>	2.3.1.4
Dernière version	<prodDate>	1.1.3.3	Langage données	<notes>	2.1.9	Segmentation	<sampProc>	2.3.1.4
Langage documentation	<notes>	1.1.6.3	Mots clés	<keyword>	2.2.1.1	Technique de collecte	<collMode>	2.3.1.6
Anonymisation	<notes>	1.1.9	Discipline	<topcClas>	2.2.1.2	Protocole de collecte	<resInstru>	2.3.1.7
Nom	<titt>	2.1.1.1	Classification	<topcClas>	2.2.1.2	Technique de collecte	<collsitu>	2.3.1.9
DDI Id	<IDNo>	2.1.1.5	Résumé	<abstract>	2.2.2	Durée des observations	<collsitu>	2.3.1.9
Affiliation	<AuthEnty>	2.1.2.1	Période couverte	<timePrd>	2.2.3.1	Analyse	<notes>	2.3.2
Auteur(s)	<AuthEnty>	2.1.2.1	Dates de récolte des données	<collDate>	2.2.3.2	Localisation archives	<accsPlac>	2.4.1.1
Copyright	<copyright>	2.1.3.2	Pays	<nation>	2.2.3.3	_____	<origArch>	2.4.1.2
Dates de l'enquête	<prodDate>	2.1.3.3	Couverture géographique	<geogcover>	2.2.3.4	Nombre de documents	<fileQnty>	2.4.1.6
Logiciel	<software>	2.1.3.5	Unités spatialisées	<geogUnit>	2.2.3.5	Engagement	<confDec>	2.4.2.1
Agence de financement	<fundAg>	2.1.3.6	Unités d'observation	<anlyUnit>	2.2.3.6	_____	<restrctn>	2.4.2.3
Bourse	<fundAg>	2.1.3.6	Niveau de comparatisme	<Universe>	2.2.3.7	Contact	<contact>	2.4.2.4
Distributeur	<distrbtr>	2.1.4.1	Groupes de population ciblés	<Universe>	2.2.3.7	_____	<citReq>	2.4.2.5
_____	<depositr>	2.1.4.3	Types de documents	<dataKind>	2.2.3.8	Conditions d'accès	<conditions>	2.4.2.7
_____	<depDate>	2.1.4.4	Transcription	<dataKind>	2.2.3.8	Publications associées	<relPubl>	2.5.3
_____	<distDate>	2.1.4.5	Périodicité	<timeMeth>	2.3.1.1	Nombre d'observations	<caseQnty>	3.1.4.1

```

1 <?xml version="1.0" encoding="utf-8"?>
2 <codeBook version="1.2.2"
3   ID="cdsp_bq_s5"
4   xml-lang="en" xmlns="http://www.icpsr.umich.edu/DDI"
5   xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
6   xsi:schemaLocation="http://www.icpsr.umich.edu/DDI http://www.icpsr.umich.edu/DDI/Version1-2-2.xsd">
7   <docDescr>
8     <citation>
9       <titlStm>
10        <titl>XML codeBook for cdsp_bq_s5</titl>
11      </titlStm>
12      <prodStm>
13        <producer affiliation="Centre de données-socio-politiques" abbr="CDSP" role="éditeur de la documentation des données">Selma Bendjal
14        <copyright>Centre de données socio-politiques (CDSP) / beQuali</copyright>
15        <prodDate date="2017-05-04">4 mai 2017</prodDate>
16      </prodStm>
17      <verStm>
18        <notes>Français</notes>
19      </verStm>
20      <notes resp="beQuali">Anonymisation réalisée par l'équipe de recherche (pseudos) et à la marge par beQuali (leaders associatifs et tabl
21    </citation>
22  </docDescr>
23  <studyDescr>
24    <citation>
25      <titlStm>
26        <titl>De l'Afrique à la France, d'une génération à l'autre</titl>
27        <IDNo agency="Centre de données socio-politique (CDSP)" ID="DOIxxx">cdsp_bq_s5</IDNo>
28      </titlStm>
29      <rspStm>
30        <AuthEnty affiliation="FACTE (CNRS/Sciences Po)" ID="ISNI0000000120201262">Jacques BAROU (coord.)</AuthEnty>
31        <AuthEnty affiliation="Caisse nationale assurance vieillesse (CNAV)" ID="IdRef15984861X">Sabrina AOUICI</AuthEnty>
32        <AuthEnty affiliation="CNAV" ID="ISNI000000007975784X">Claudine ATTIAS-DONFUT</AuthEnty>
33        <AuthEnty affiliation="CNAV" ID="IdRef057668507">Rémi GALLOU</AuthEnty>
34        <AuthEnty affiliation="Institut interdisciplinaire d'anthropologie du contemporain (IIAC)" ID="ISNI0000000108667086">Nicole LAPIERRI
35        <AuthEnty affiliation="CNAV" ID="ISNI00000000053760653">Alain ROZENKIER</AuthEnty>
36        <AuthEnty affiliation="Université Paris X Nanterre" ID="ISNI0000000109169476">Martine SEGALIN</AuthEnty>
37      </rspStm>
38    </prodStm>
39    <copyright>CDSP/beQuali, Crédits photos : Jacques Barou - Tetraktys ; Sabrina Aouici - copyright ; Rémi Gallou - copyright ; Claudir
40    <prodDate date="2006-00-00">2006</prodDate>
41    <prodDate date="2011-00-00">2011</prodDate>
42    <software>PowerGrep</software>
43    <software>Exmaralda</software>
44    <fundAg abbr="CNAV" role="financier" ID="ISNI0000000121902300">Caisse nationale d'assurance vieillesse (CNAV)</fundAg>
45    <fundAg role="Bourse">Aucune</fundAg>

```

```

46 </proStat>
47 </distStm>
48 <distrbtr abbr="CDSF" URI="http://www.bequali.fr" ID="ISNI0000000404630798">Centre de données socio-politiques / beQuali</distrbtr>
49 <distrbtr URI="https://quetelet.casd.eu/fr/">Réseau Quetelet</distrbtr>
50 <depositr affiliation="PACTE">BAROU Jacques</depositr>
51 <depositr affiliation="CNAV">GALLOU Rémi</depositr>
52 <depositr affiliation="CNAV">AOUICI Sabrina</depositr>
53 <depDate date="2016-02-19">19 février 2016</depDate>
54 <distDate date="2017-05-20">20 mai 2017</distDate>
55 </distStm>
56 <verStm>
57 <version>facultatif (en cas de deuxième version de publication de l'enquête sur beQuali)</version>
58 </verStm>
59 <biblCit>MICHELAY Guy, SIMON Michel, DONÉGANI Jean-Marie [auteurs] : Représentation du champ social, attitudes politiques et changements :
60 <notes>Français, Anglais</notes>
61 </citation>
62 <stdyInfo>
63 <subject>
64 <keyword vocab="site CDSF">Immigration</keyword>
65 <keyword vocab="RAMEAU">Immigrés</keyword>
66 <keyword vocab="RAMEAU">Enfants d'immigrés</keyword>
67 <keyword vocab="site CDSF">Famille</keyword>
68 <keyword vocab="RAMEAU">Parents et enfants</keyword>
69 <keyword vocab="RAMEAU">Intégration</keyword>
70 <keyword vocab="ELSST">>Génération (âge)</keyword>
71 <keyword vocab="RAMEAU">Afrique</keyword>
72 <keyword vocab="RAMEAU">France</keyword>
73 <keyword vocab="Y">2006</keyword>
74 <keyword vocab="Y">2010</keyword>
75 <keyword/>
76 <topcClas vocab="RAMEAU">Sociologie</topcClas>
77 <topcClas>Sociologie de l'immigration</topcClas>
78 <topcClas>Sociologie de la famille</topcClas>
79 <topcClas>Ethnographie</topcClas>
80 </subject>
81 <abstract>Enquête ethnographique portant sur les immigrés d'origine subsaharienne, réalisée entre 2006 et 2011 (en France) par Jacques Ba.
82 L'enquête interroge le statut des immigrés noirs en France, à travers leurs trajectoires et les représentations qu'ils s'en font.
83 La comparaison intergénérationnelle des visions de la migration, du pays et de la culture d'origine, ou encore de la famille apporte un é.
84 Nota bene : deux enquêtes ont été menées en parallèle au Royaume-Uni et en Afrique du Sud. L'équipe de recherche française a notamment co.
85 </abstract>
86 <sumDescr>
87 <timePrd date="2006-00-00" event="start">2006</timePrd>
88 <timePrd date="2010-00-00" event="end">2010</timePrd>
89 <collDate date="2006-00-00" event="start">2006</collDate>
90 <collDate date="2010-00-00" event="end">2010</collDate>

```

```

91 <nation>France</nation>
92 <nation>Afrique sub-saharienne (Angola, Bénin, Burkina Faso, Cameroun, Centrafrique, Comores, Congo Brazzaville, Côte d'Ivoire, Gabon, Ghana, Guinée, Mali, Niger, République
93 <geogCover>Monde</geogCover>
94 <geogUnit>Région parisienne</geogUnit>
95 <geogUnit>Région lyonnaise</geogUnit>
96 <anlyUnit>Famille</anlyUnit>
97 <anlyUnit>Individu</anlyUnit>
98 <universe level="study">Comparaison internationale et générationnelle</universe>
99 <universe level="study">Familles d'origine d'Afrique subsaharienne ayant vécu une immigration en France : parents et enfants à une génération d'écart, Leaders associatifs
100 <dataKind>Documents textuels, enregistrements audio</dataKind>
101 <dataKind>Transcription verbale et paraverbale simple, synthèses familiales saisies a posteriori, notes d'observation.</dataKind>
102
103 </sumDescr>
104 </studyInfo>
105 <method>
106 <dataColl>
107 <timeMeth>Entretiens uniques</timeMeth>
108 <sampProc>Accès aux enquêtés par réseau de connaissances personnelles (associatif, etc.)</sampProc>
109 <sampProc>Constitution de l'échantillon par pays d'origine (Angola, Bénin, Burkina Faso, Cameroun, Centrafrique, Comores, Congo Brazzaville, Côte d'Ivoire, Gabon, Ghana, République
110 <sampProc>Segmentation par âge et sexe pour les focus groups (mères, pères, filles, fils) ; par âge (parents, enfants) pour les entretiens familiaux.</sampProc>
111 <collMode>Entretiens individuels semi-directifs, focus groups animés par l'équipe de recherche</collMode>
112 <resInstru>Guides d'entretien semi-directif pour les entretiens individuels et les focus groups (4 guides) élaborés en collaboration avec l'équipe anglaise ; questionnaire
113 <collSitu>9 enquêteurs, 22 familles interrogées, 65 enquêtés au total (36 femmes, 27 hommes) dont 19 participants aux focus groups (5 pères, 6 mères, 5 fils, 4 filles), 1
114 </dataColl>
115 <notes>Analyse interprétative simple</notes>
116 </method>
117 <dataAccs>
118 <setAvail>
119 <accsPlac URI="http://bequali.fr/fr/les-enquetes/lenquete-en-bref/cdsp_bq_s5">www.bequali.fr</accsPlac>
120 <accsPlac URI="https://www.pacte-grenoble.fr/content/accueil">PACTE</accsPlac>
121 <accsPlac URI="http://www.statistiques-recherches.cnav.fr">CNAV</accsPlac>
122 <origArch>PACTE, CNAV</origArch>
123 <fileQnty>219 fichiers (dont 157 à disposition)</fileQnty>
124 </setAvail>
125 <useStmt>
126 <confDec formNo="1">Pour accéder aux données, l'utilisateur doit signer un engagement de confidentialité.</confDec>
127 <restrctn>En préparant les fichiers de données de cette enquête, les auteurs ont utilisé des pseudonymes, et le CDSP a retiré les identifiants directs et caractéristiques
128 <contact affiliation="CDSP" URI="http://www.bequali.fr">contact@bequali.fr</contact>
129 <citReq>Les publications basées sur les enquêtes diffusées par le CDSP doivent citer leurs sources par des références bibliographiques. Afin de s'assurer que l'attribution
130 <conditions>Accès soumis à autorisation</conditions>
131 </useStmt>
132 </dataAccs>
133 <otherStudyWat>
134 <relPubl>Aouici Sabrina et Rémi Gallou, « Ancrage et mobilité de familles d'origine africaine : regards croisés de deux générations », Revue internationale Enfances, Familles
135 <relPubl>Barou Jacques (dir.), De l'Afrique à la France. D'une génération à l'autre, Paris : Armand Colin, Collection Sociétales, 2011, 256 p. ISBN : 9782200255848.</relPubl>

```

Décrire un corpus d'archives : le standard EAD

Encoded Archival Description

Language XML hiérarchique → Vision et description arborescente (niveaux dossiers, sous-dossiers, fichiers)

Description archivistique → Interopérabilité (Calames.abes.fr : catalogue des archives de l'ESR)

Champs principaux

- Cote, Intitulé, Dates extrêmes, Langue
- Description matérielle
- Classement
- Présentation du contenu
- Bibliographie
- Organisme responsable de l'accès
- Informations sur les modalités d'entrée
- Historique de la conservation
- Localisation physique
- Conditions d'accès et d'utilisation

```

4 <eadheader>
5 <eadid>A000_29662</eadid>
6 <filedesc>
7 <titlestmt>
8 <titleproper>Des femmes en politique</titleproper>
9 </titlestmt>
10 <publicationstmt>
11 <publisher>Archive</publisher>
12 <address/>
13 </publicationstmt>
14 </filedesc>
15 <profiledesc>
16 <creation>Cet instrument de recherche a été généré avec la suite Ligeo Archives<date normal="2017/2017">16/06/2017</date></creation>
17 <langusage>Instrument de recherche rédigé en<language>français</language></langusage>
18 </profiledesc>
19 </eadheader>
20 <archdesc id="ligeo-29662">
21 <did>
22 <unittitle>Des femmes en politique</unittitle>
23 <unitid>BO sp6</unitid>
24 <physdesc>
25 <extent unit="ml" type="longueur">0.6</extent>
26 </physdesc>
27 <unitdate normal="1984/1985" label="ligeo-date-automatique">1984-1985</unitdate>
28 </did>
29 <custodhist>
30 <p>2016-09-01</p>
31 <p>Archives d'enquêtes traitée par le CDSP dans le cadre de leur mise à disposition sur beQuali.</p>
32 </custodhist>
33 <physloc>Bâtiment Ligeo &gt; Rez-de-chaussée &gt; Salle des entrées</physloc>
34 <arrangement>
35 <note>
36 <p>Plan de classement</p>
37 </note>
38 </arrangement>
39 <scopecontent>
40 <p>Enquête réalisée entre 1984 et 1985 (en île de France) par Mariette Sineau dans le cadre d'une recherche au Centre de recherches politiques de Sc
41 </scopecontent>
42 <origination>
43 <corpname authfilenumber="716">Centre de recherche politique de Sciences Po (CEVIPOF)</corpname>
44 <persname authfilenumber="18775">Sineau, Mariette</persname>
45 </origination>
46 <nhvstech type="etat_materiel">A</nhvstech>
47 <accessrestrict>
48 <legalstatus>Archives publiques</legalstatus>
49 </accessrestrict>
50 <acqinfo>
51 <chronlist>

```

```

58 <dsc type="in-depth">
59   <c altrender="ligeo-branche-standard" id="ligeo-29821">
60     <did>
61       <unittitle prep</unittitle>
62       <unitid type="ligeo-cote-automatique">BQ sp6 1 à 10</unitid>
63       <unitdate label="ligeo-date-automatique" normal="1984">1984</unitdate>
64     </did>
65   <c altrender="ligeo-branche-standard" id="ligeo-29822">
66     <did>
67       <unittitle admi</unittitle>
68       <unitid type="ligeo-cote-automatique">BQ sp6 1 à 2</unitid>
69     </did>
70   <c altrender="ligeo-article-standard" id="ligeo-29823" level="item">
71     <did>
72       <unittitle cdsp_bequali_sp6_prep_admi_corresp_FR_18_lettre</unittitle>
73       <unitid>BQ sp6 1</unitid>
74       <unitid type="ancienne_cote">23 ENQ 2</unitid>
75       <unitdate label="ligeo-date-automatique">1984-10-23</unitdate>
76     </did>
77     <scopecontent>
78       <p>Lettre de Madame la Maire Paulette Fost à Mariette Sineau : réponse à la sollicitation d'interview</p>
79     </scopecontent>
80   </c>
81   <c altrender="ligeo-article-standard" id="ligeo-29824" level="item">
82     <did>
83       <unittitle cdsp_bequali_sp6_prep_admi_corresp_FR_20_lettre</unittitle>
84       <unitid>BQ sp6 2</unitid>
85       <unitid type="ancienne_cote">23 ENQ 2</unitid>
86       <unitdate label="ligeo-date-automatique">1984-10-18</unitdate>
87     </did>
88     <scopecontent>
89       <p>Lettre de Madame l'adjoite au Maire Nicole Chouraoui à Mariette Sineau : réponse à la sollicitation d'interview</p>
90     </scopecontent>
91   </c>
92 </c>
93 <c altrender="ligeo-branche-standard" id="ligeo-29825">
94   <did>
95     <unittitle docu</unittitle>
96     <unitid type="ligeo-cote-automatique">BQ sp6 3 à 10</unitid>

```

1 : Sous-fonds

2 : Dossier

3 : Fichier

La visualisation, un autre outil de contextualisation

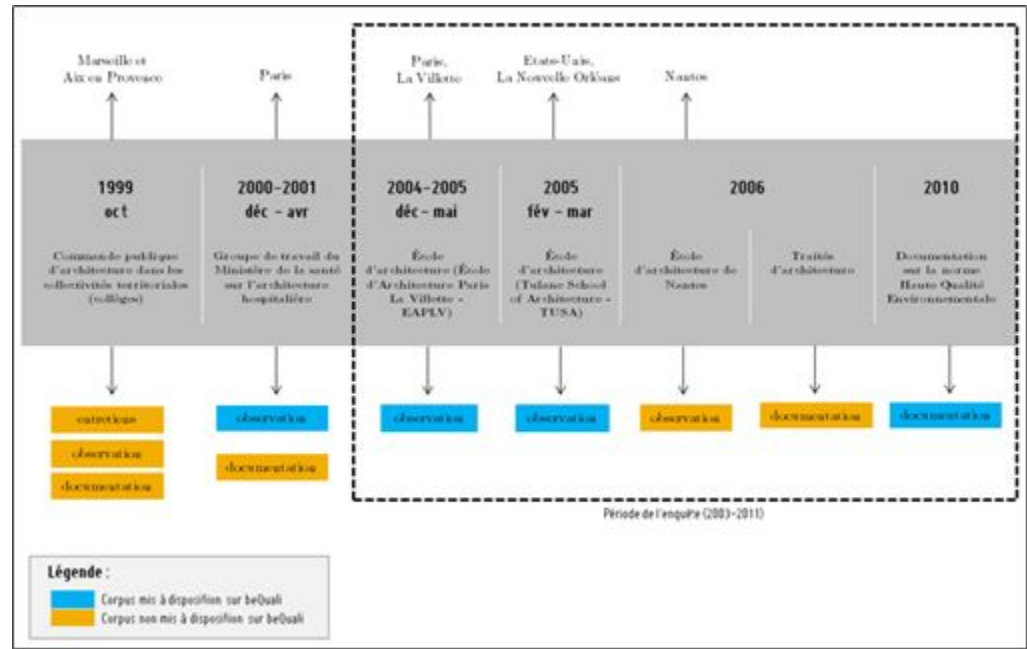
Des frises chronologiques
et des cartes heuristiques
pour décrire des corpus complexes

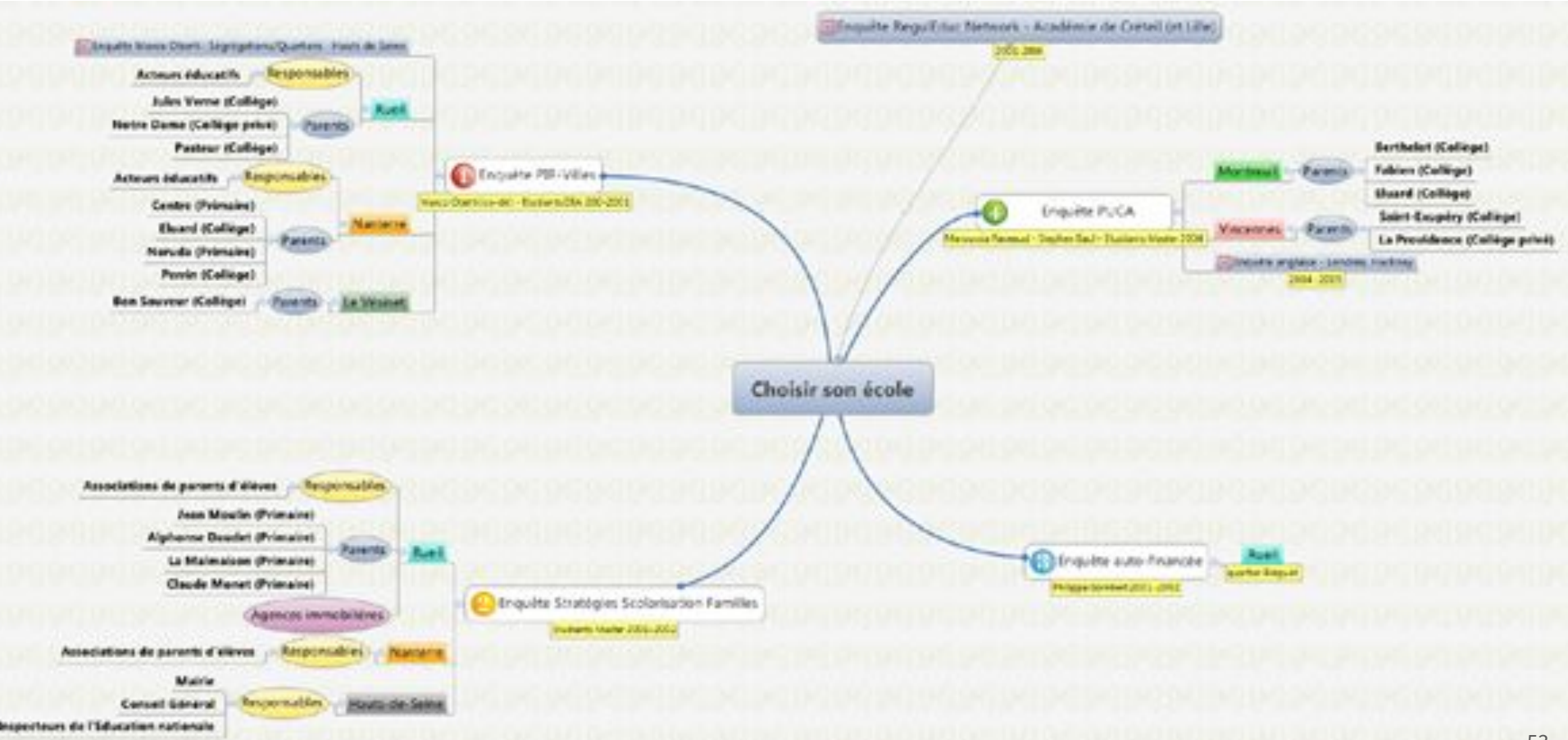
3 exemples :

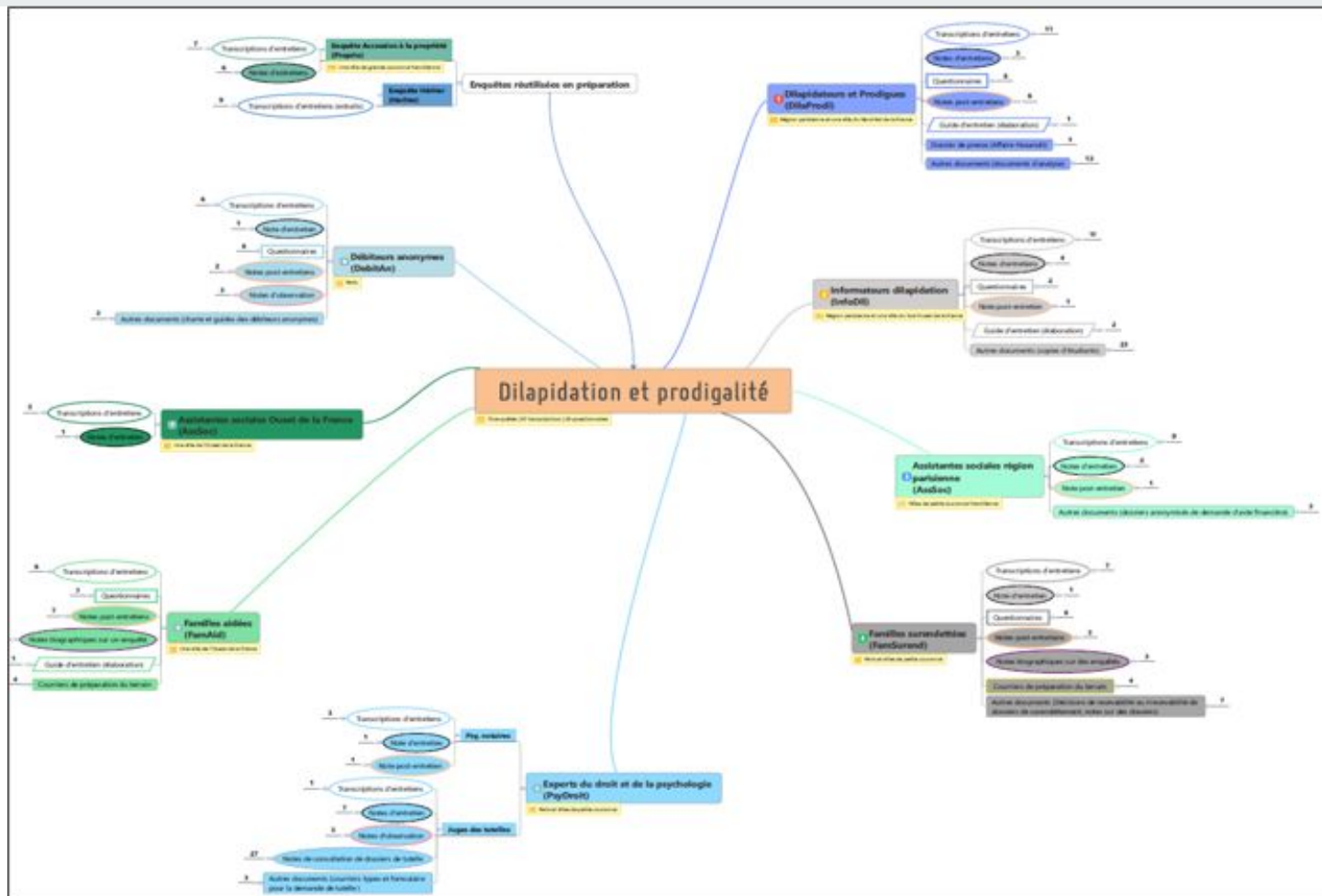
Florent Champy, *La profession d'architecte*

Agnès van Zanten, *Choisir son école*

Anne Gotman, *Dilapidation et prodigalité*







beQuali



Explorer le corpus via la localisation géographique ⓘ

Inclure aux résultats les fichiers non localisés

Genèse

de l'enquête

Une enquête initiée par la CNAV (1'03)

▶ ● 0:00 / 0:00

Une équipe aux profils de recherche différents, mais c

▶ ● 0:00 / 0:00

Une équipe de recherche internationale (0'49)

▶ ● 0:00 / 0:00

Une enquête entre rupture et continuité (0'49)

Umap

Telemme fonds Vovelle

Filter:

- Comune commençant
- VOV21-52-1-A
- VOV21-51-1-A
- VOV21-50-1-A
- VOV21-49-1-A
- VOV21-48-1-C
- VOV21-48-2-C
- VOV21-48-1-A
- VOV21-47-1-A
- VOV21-45-1-A
- VOV21-44-1-A
- VOV21-43-1-A
- VOV21-41-2-A
- VOV21-40-1-A
- VOV21-39-1-A
- VOV21-38-1-A
- VOV21-37-1-A
- VOV21-36-1-A
- VOV21-35-1-A
- VOV21-34-1-A

VoyantTools

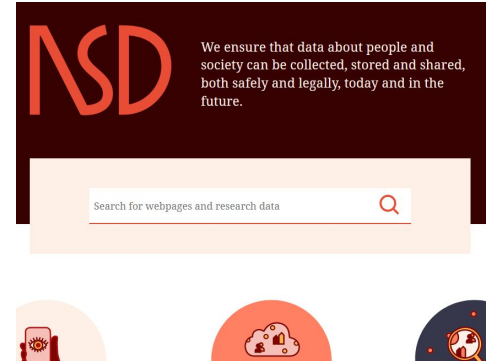


Nesstar Publisher

Développé par le NSD : Norwegian Center for Research Data

<http://www.nesstar.com/software/publisher.html>

Un outil un petit peu vieillissant... mais bien pratique pour documenter au niveau des variables.



Nesstar
About Nesstar
Nesstar 4.0 is now available
About
Customers
Privacy Policy
Contact
Software
Publisher
Server
Webview
Public API
REST API
OAI PMH Server

Nesstar Publisher

Advanced Data Management

Nesstar Publisher is an advanced data management program. It consists of data and metadata conversion and editing tools, enabling the user to prepare these materials for publication to a Nesstar Server. However, it can also be used as a stand-alone tool for the preparation of data and metadata. The Publisher enables users to enhance datasets by combining a wide range of catalogue and contextual information, which can then be viewed within the Nesstar web client, Nesstar WebView.

Main features

- The easy editing/creation of DDI documented datasets with no need to know XML



Variables

Number	Name	Width	Label	StartCol	EndCol	Record	Decimals
v7	csoc_Q01_06		Connait : un avocat	*	*	1	0
v8	csoc_Q01_07		Connait : un mécanicien automobile	*	*	1	0
v9	csoc_Q01_08		Connait : un infirmier	*	*	1	0
v10	csoc_Q01_09		Connait : un policier	*	*	1	0
v11	csoc_Q01_10		Connait : un enseignant	*	*	1	0
v12	csoc_Q02_01		Opinion : les différences de revenus en France sont trop grandes	*	*	1	0
v13	csoc_Q02_02		Opinion : pour qu'une société soit juste, les différences de niveau	*	*	1	0
v14	csoc_Q02_03		Opinion : c'est la responsabilité du gouvernement de réduire les ir	*	*	1	0
v15	csoc_Q02_04		Opinion : en France, les aides sociales rendent les gens paresseu	*	*	1	0
v16	csoc_Q03		Organisme en charge d'assurer en priorité les soins de santé	*	*	1	0
v17	csoc_Q04		Organisme en charge d'assurer la prise en charge des personnes	*	*	1	0
v18	csoc_Q05		Opinion : les gens comme moi n'ont pas leur mot à dire sur l'action	*	*	1	0
v19	csoc_Q06_01		Personne sollicitée en premier pour : l'aider dans une tâche ména	*	*	1	0
v20	csoc_Q06_02		Personne sollicitée en premier pour : l'aider à la maison si il est ma	*	*	1	0
v21	csoc_Q06_03		Personne sollicitée en premier pour : être à ses cotés si il se sent	*	*	1	0
v22	csoc_Q06_04		Personne sollicitée en premier pour : donner un conseil sur des pr	*	*	1	0
v23	csoc_Q06_05		Personne sollicitée en premier pour : partager un moment convivi	*	*	1	0
v24	csoc_Q07_01		Personne sollicitée en premier pour : emprunter une grosse somm	*	*	1	0
v25	csoc_Q07_02		Personne sollicitée en premier pour : trouver un travail	*	*	1	0
v26	csoc_Q07_03		Personne sollicitée en premier pour : des problèmes administratifs	*	*	1	0
v27	csoc_Q07_04		Personne sollicitée en premier pour : trouver un endroit où loger	*	*	1	0

Category Hierarchy

- 1 - Tout à fait d'acc
- 2 - Plutôt d'accord
- 3 - Ni d'accord, ni p
- 4 - Plutôt pas d'acc
- 5 - Pas du tout d'ac
- 6 - Ne peut choisir
- 9 - N'a pas répondu

Value: Label:

Category Text:

Level Name:

GeoMap URI:

Variable information

Data Type:

Measure:

Is Time Variable

Is Weight Variable

Min: Max: Decimals:

Implicit decimals

Missing data:

Documentation

Statistics

Include Weighted Statistics

Include Frequencies

List Missing At End

Sorting of Frequencies:

Summary Statistics Options:

Include Valid

Include Min

Include Max

Include Mean

Include Weighted Mean

Include StdDev

Include Weighted StdDev

Weights | Documentation

Frequencies:

Value	Label	N		
1	Tout à fait d'accord	412	<div style="width: 23.1%;"></div>	23.1%
2	Plutôt d'accord	652	<div style="width: 36.6%;"></div>	36.6%
3	Ni d'accord, ni pas d'accord	374	<div style="width: 21%;"></div>	21%
4	Plutôt pas d'accord	251	<div style="width: 14.1%;"></div>	14.1%
5	Pas du tout d'accord	63	<div style="width: 3.5%;"></div>	3.5%
6	Ne peut choisir	29	<div style="width: 1.6%;"></div>	1.6%
9	N'a pas répondu	0		Missing

Summary Statistics:

Type	Value
Valid	1781

Variables

Number	Name	Width	Label	StartCol	EndCol	Record	Decimals
v7	csoc_Q01_06		Connait : un avocat	*	*	1	0
v8	csoc_Q01_07		Connait : un mécanicien automobile	*	*	1	0
v9	csoc_Q01_08		Connait : un infirmier	*	*	1	0
v10	csoc_Q01_09		Connait : un policier	*	*	1	0
v11	csoc_Q01_10		Connait : un enseignant	*	*	1	0
v12	csoc_Q02_01		Opinion : les différences de revenus en France sont trop grandes	*	*	1	0
v13	csoc_Q02_02		Opinion : pour qu'une société soit juste, les différences de niveau	*	*	1	0
v14	csoc_Q02_03		Opinion : c'est la responsabilité du gouvernement de réduire les in	*	*	1	0
v15	csoc_Q02_04		Opinion : en France, les aides sociales rendent les gens paresseu	*	*	1	0
v16	csoc_Q03		Organisme en charge d'assurer en priorité les soins de santé	*	*	1	0
v17	csoc_Q04		Organisme en charge d'assurer la prise en charge des personnes	*	*	1	0
v18	csoc_Q05		Opinion : les gens comme moi n'ont pas leur mot à dire sur l'action	*	*	1	0
v19	csoc_Q06_01		Personne sollicitée en premier pour : l'aider dans une tâche ména	*	*	1	0
v20	csoc_Q06_02		Personne sollicitée en premier pour : l'aider à la maison si il est ma	*	*	1	0
v21	csoc_Q06_03		Personne sollicitée en premier pour : être à ses cotés si il se sent	*	*	1	0
v22	csoc_Q06_04		Personne sollicitée en premier pour : donner un conseil sur des pr	*	*	1	0
v23	csoc_Q06_05		Personne sollicitée en premier pour : partager un moment convivi	*	*	1	0
v24	csoc_Q07_01		Personne sollicitée en premier pour : emprunter une grosse somm	*	*	1	0
v25	csoc_Q07_02		Personne sollicitée en premier pour : trouver un travail	*	*	1	0
v26	csoc_Q07_03		Personne sollicitée en premier pour : des problèmes administratifs	*	*	1	0
v27	csoc_Q07_04		Personne sollicitée en premier pour : trouver un endroit où louer	*	*	1	0

▼ Documentation

Statistics

Weights

Documentation

- 4.2 Variable Descriptions - Questions
 - Pre-Question Text 4.2.8.1
 - Literal Question 4.2.8.2
 - Post-Question Text 4.2.8.3
 - Interviewer Instructions 4.2.8.6
- 4.2 Variable Descriptions
 - Universe 4.2.12
 - Total Responses 4.2.13
 - Variable Text 4.2.15
- 4.2 Variable Descriptions - Concepts
 - Concepts 4.2.21
- 4.2 Variable Descriptions - Notes
 - Notes 4.2.24

Pre-Question Text 4.2.8.1

Literal Question 4.2.8.2

Êtes-vous d'accord ou pas d'accord avec la proposition suivante :
 - C'est la responsabilité du gouvernement de réduire les différences entre les revenus importants et les revenus faibles

Category Hierarchy

- 1 - Tout à fait d'accord
- 2 - Plutôt d'accord
- 3 - Ni d'accord, ni pas d'accord
- 4 - Plutôt pas d'accord
- 5 - Pas du tout d'accord
- 6 - Ne peut choisir
- 9 - N'a pas répondu

Value: Label:

Category Text:

Level Name:

GeoMap URI:

▼ Variable information

Data Type:

Numeric

Measure:

Nominal

 Is Time Variable Is Weight Variable

Min:

1

Max:

6

Decimals:

0

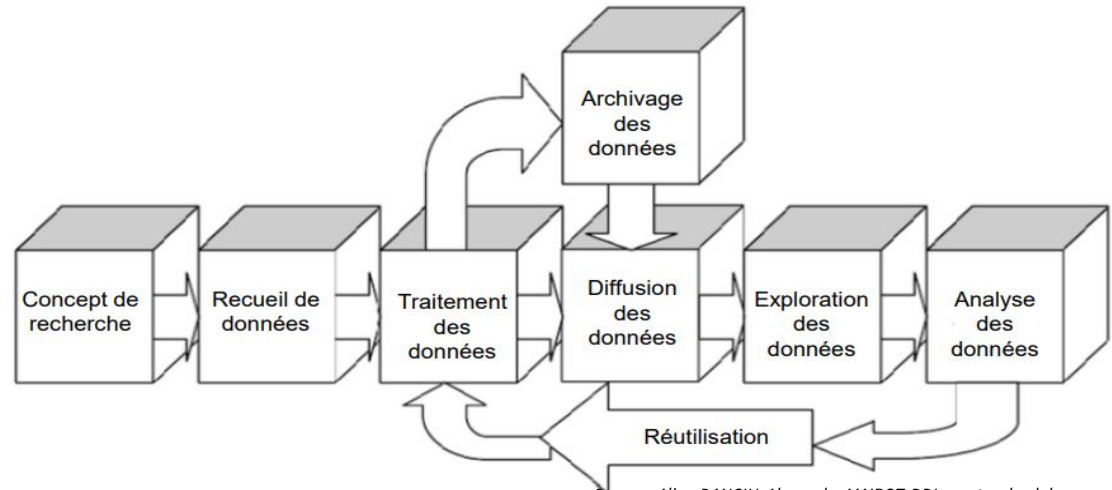
 Implicit decimals

Missing data:

>= 9

*

Colectica au service de DDI-Lifecycle



Source : Alina DANCIU, Alexandre MAIROT, DDI, un standard de documentation des données, Tutoriels MATE-SHS, 14 mars 2019
<https://mate-shs.cnrs.fr/actions/tutomate/tuto20-danciu-mairot-ddi/>

Un exemple

“Je souhaite analyser l'évolution de l'intérêt pour la politique au moment des élections présidentielles de 2007 grâce aux enquêtes du Baromètre Politique Français.”



D'une documentation éparpillée ...

Enquête Vague 1

*Base de données
Questionnaire
Livre de codes*

Enquête Vague 2

*Base de données
Questionnaire
Livre de codes*

Enquête Vague 3

*Base de données
Questionnaire
Livre de codes*

Enquête Vague 4

*Base de données
Questionnaire
Livre de codes*

Q47: Personnalité dont on se sent le plus proche

Information	[Series Name= continuous] [Format=numeric] [Range= 1-18] [Missing=*]
Statistics [NW/W]	[Valid=5650 / 5650.000] [[Invalid=0 / 0]
Littéral question	El parmi toutes ces personnalités que je viens de vous citer, quelle est celle dont vous vous sentez le plus proche ?
Interviewer's instructions	CITER - UNE SEULE REPONSE POSSIBLE UNE RELANCE ET UNE SEULE SUR LE NSP
Notes	Rotation des items

Value	Label	Cases	Weighted	Percentage (Weighted)
1	Bayrou	275	349.7	6.2%
2	Besanconot	292	283.1	5.0%
3	Bové	212	243.4	4.3%
4	Buffet	143	130.7	2.3%
5	Chirac	321	297.4	5.3%
6	Fabius	129	122.4	2.2%
7	Hollande	91	66.9	1.2%
8	Jospin	702	523.2	9.3%
9	Laguiller			
10	Le Pen			
11	Royal			
12	Sarkozy			

Q47: Personnalité dont on se sent le plus proche

Information	[Series Name= continuous] [Format=numeric] [Range= 1-20] [Missing=*]
Statistics [NW/W]	[Valid=5240 / -] [[Invalid=0 / -]
Littéral question	Parmi les personnalités suivantes, quelle est celle dont vous vous sentez le plus proche ?
Interviewer's instructions	CITER - UNE SEULE REPONSE POSSIBLE UNE RELANCE ET UNE SEULE SUR LE NSP
Notes	Rotation des items sauf "aucune d'entre elles"

Value	Label	Cases	Percentage
1	François Bayrou	438	8.4%
2	Olivier Besanconot	327	6.2%
4	Mario-George Buffet	138	2.6%
5	Jacques Chirac	326	6.2%
9	Arlette Laguiller	150	2.9%
10	Jean-Marie Le Pen	171	3.3%
11	Ségolène Royal	1515	28.9%
12	Nicolas Sarkozy	1310	25.0%
14	Dominique De Villepin	101	1.9%
15	Philippe De Villiers	98	1.9%
16	Dominique Voynet	161	3.1%
17	(Aucune d'entre elles)	232	4.4%
18	(Nsp)	24	0.5%
19	Michèle Alliot-Marie	192	3.7%
20	Jean-Pierre Chevènement	57	1.1%

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

Q47: Personnalité dont on se sent le plus proche

Information	[Series Name= continuous] [Format=numeric] [Range= 1-18] [Missing=*]
Statistics [NW/W]	[Valid=5647 / 5647.008] [[Invalid=0 / 0]
Littéral question	Parmi les personnalités suivantes, quelle est celle dont vous vous sentez le plus proche ?
Interviewer's instructions	CITER - UNE SEULE REPONSE POSSIBLE UNE RELANCE ET UNE SEULE SUR LE NSP

- 89 -

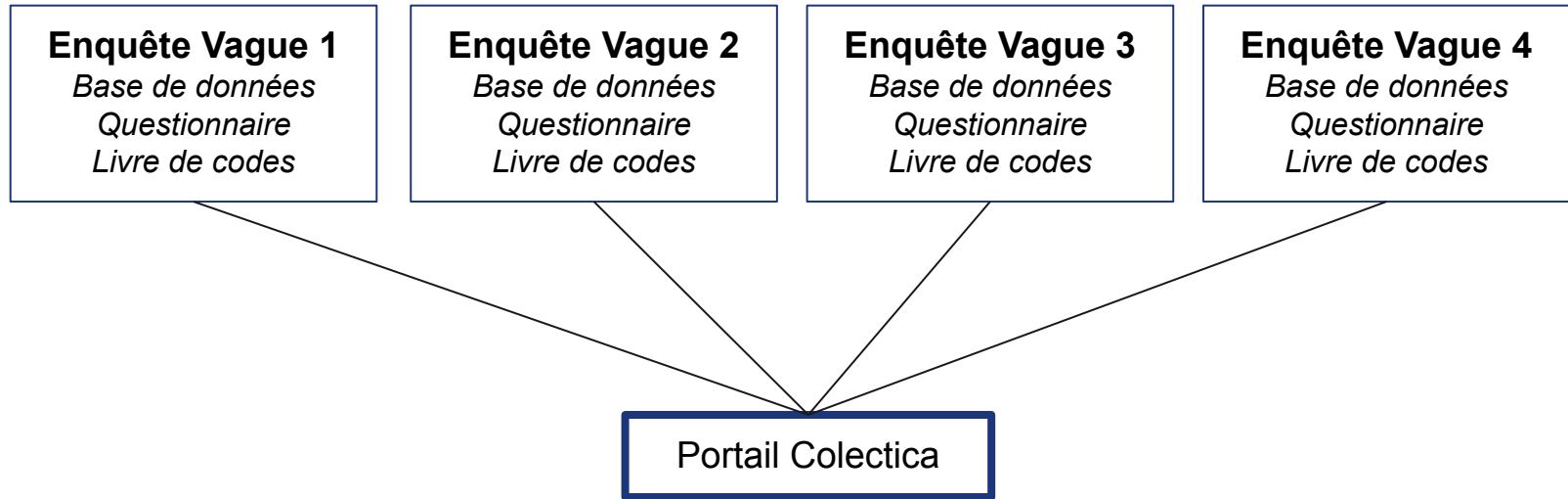
Q47: Personnalité dont on se sent le plus proche

Information	[Series Name= continuous] [Format=numeric] [Range= 1-18] [Missing=*]
Statistics [NW/W]	[Valid=5239 / -] [[Invalid=0 / -]
Littéral question	Parmi les personnalités suivantes, quelle est celle dont vous vous sentez le plus proche ?
Interviewer's instructions	CITER - UNE SEULE REPONSE POSSIBLE UNE RELANCE ET UNE SEULE SUR LE NSP
Notes	Rotation des items sauf "aucune d'entre elles"

Value	Label	Cases	Percentage
1	François Bayrou	808	15.4%
2	Olivier Besanconot	326	6.2%
3	José Bové	227	4.3%
4	Mario-George Buffet	176	3.4%
5	Jacques Chirac	350	6.7%
9	Arlette Laguiller	180	3.4%
10	Jean-Marie Le Pen	149	2.8%
11	Ségolène Royal	1159	22.1%
12	Nicolas Sarkozy	1308	25.0%
14	Dominique De Villepin	132	2.5%
15	Philippe De Villiers	133	2.5%
16	Dominique Voynet	107	2.0%
17	(Aucune d'entre elles)	161	3.1%
18	(Nsp)	23	0.4%

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

... à un point d'accès unique.



L'exploration via le portail Colectica

- + QUEST**
Identifiant
- + VAGUE**
Vague d'enquête
- + DATE_TER**
Date de terrain
- + DPT**
Département
- + REG**
Région
- + CCM**
Catégorie de commune selon espace Urbain / Rural
- + AGGLO**
Catégorie d'agglomération
- + GR**
Grande région
- + REGA**
Région administrative
- + NATIO**
Nationalité
- + INSCR**
Inscription sur listes électorales
- + SEXE**
Sexe
- + ANNEE**
Année de naissance
- + RAGE**
Tranche d'âge
- + AGE**
Age
- + RS1**
Age de fin d'études
- + RS2B**
Catégorie de diplôme obtenu

< ALEA4
BPF2007-R1 (25 of 236)
Q4 >

Variable Description

Name Q3
Label Intérêt pour la politique
Dataset bpr1

Value	Label	Frequency	% of valid	% of all
1	Beaucoup	674	11.93%	11.93%
2	Assez	1,843	32.62%	32.62%
3	Peu	2,213	39.17%	39.17%
4	Pas du tout	918	16.25%	16.25%
5	NSP	2		0.04%

Valid	Invalid	Min	Max
5650	0	1	5

Representation

Type Code List
Selection Style SelectOne
Codes opinion.scale00001

- 1 Beaucoup
- 2 Assez
- 3 Peu
- 4 Pas du tout

Role input
Aggregation Method Unspecified
Temporal False
Geographic False
Represented Variable **+** Intérêt pour la politique

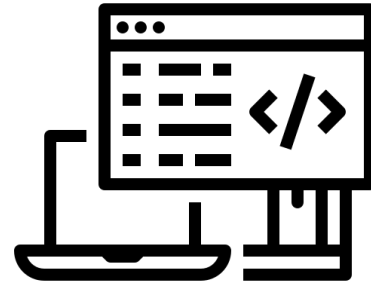
L'expérimentation Colectica au CDSP

- Un projet : ANR Flash Science Ouverte - *UpMet WP2*
- Des outils de la suite Colectica: Designer, Repository, Portal
- Un objectif : la création d'une base de métadonnées

Bilan de l'expérience

- Usages concluants : gestion, harmonisation, exploration des données **longitudinales**
- Usage moins pertinent : enquêtes “uniques”
- Investissement initial important (formation, infrastructure)

Des questions ?



Colectica Designer - Fonctionnalités

Création ...

... des métadonnées.

Import

... des jeux de données, questionnaires, DDI, listes de métadonnées.

Ré-utilisation ...

... facilitée par un référencement unique pour chaque objets de la base.

Connexion ...

... des éléments pour construire la documentation et créer des liens.

La confidentialité des données au défi du partage de données en auto-dépôt

Contexte

- Le contexte d'une expérimentation de l'auto-dépôt de données en SHS, dans le cadre du projet UpMET
- Les expériences de plus de 15 ans du CDSP en documentation d'enquêtes quanti et quali

L'auto-dépôt de données 🚗

- Exemple Zenodo ou Nakala.
- Les équipes de recherche seules en selle 🚲.
- Expérimentation UpMet : un accompagnement et une validation.

Les intérêts

Pour les équipes de recherche :

- Être maître de leur temps et de leurs données.
- Permettre à plus d'équipes de partager leurs données.

Pour le CDSP :

- Pouvoir augmenter le nombre de données diffusées et mettre en place de nouveaux types de collaborations.

Les inconvénients

- Augmentation de la charge de travail sur les équipes de recherche.
- Risque de l'existence de données faiblement documentées, mal nommées, peu compréhensibles et peu réutilisables, bref peu FAIR.
- Risque de rupture de la confidentialité.

La plateforme data.sciencespo

SciencesPo

Ajouter des données ▾ Rechercher ▾ Guide d'utilisation Support Français ▾ S'inscrire Se connecter

data.sciencespo

📊 Statistiques

3 787 téléchargements

✉ Contact 🔄 Partager

L'entrepôt institutionnel de données de la recherche de Sciences Po

Chercher dans ce dataverse...

🔍 Trouver

Recherche avancée

👤 Dataverses (16)

📁 Jeux de données (336)

📄 Fichiers (1 149)

Dataverse Category

Laboratoire (13)

Research Project (2)

Organization or Institution (1)

Metadata Source

data.sciencespo (336)

Harvested (16)

Publication Year

2020 (328)

2021 (8)

Classification des sujets Terme

Élections (172)

1 à 10 de 352 résultats

📄 Trier ▾

Pratiques numériques - vague 1 (2013)

11 oct. 2021 - ELIPSS

Équipe ELIPSS, 2020, "Pratiques numériques - vague 1 (2013)", <https://doi.org/10.21410/7E4/DWHY6T>, data.sciencespo, V3

L'enquête "Pratiques numériques" (PN) mesure l'évolution des comportements liés aux technologies de l'information et de la communication (TIC) dans un panel où les répondants ont été équipés d'une tablette tactile connectée. L'enquête est répétée annuellement. S'inscrivant dans l...

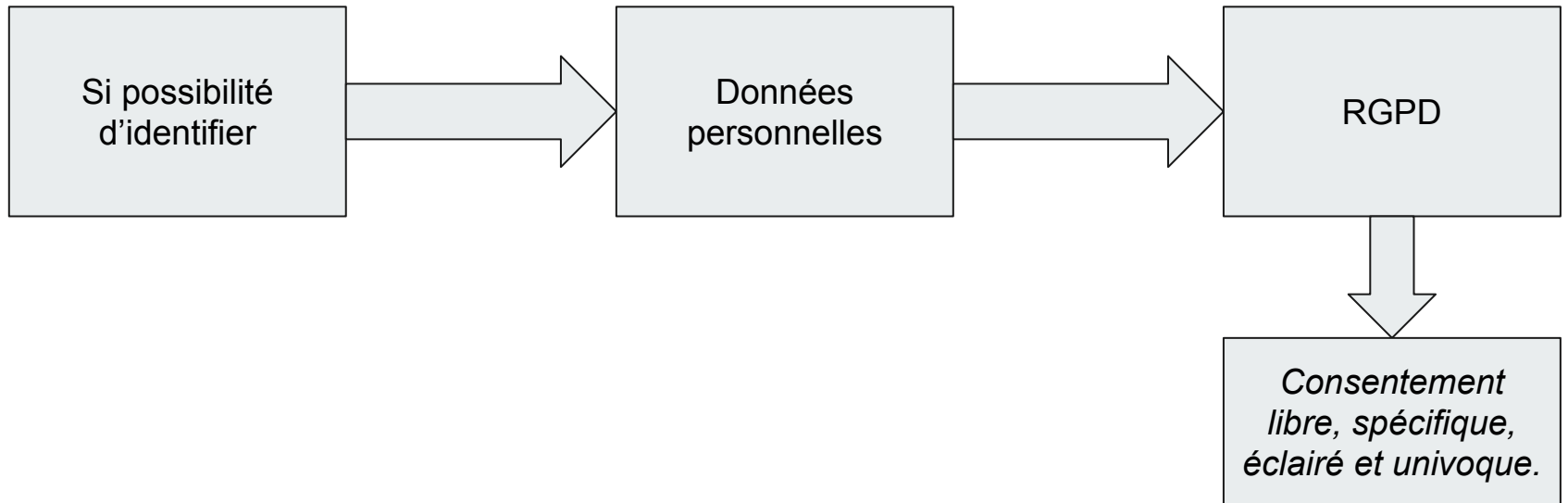
Enquête annuelle - vague 1 (2013)

11 oct. 2021 - ELIPSS

Le défi de l'auto-dépôt.

Transmettre les **bonnes pratiques** aux équipes de recherches.

Un rappel sur les données personnelles et le RGPD



Partager des données d'enquêtes :

“As open as possible, as closed as necessary”

Quatres possibilités pour le partages des données avec données personnelles:

- Formulaires de consentement
- Anonymiser les données
- Comité d'éthique
- Ne pas partager les données

Anonymiser :

*“L’anonymisation est un traitement qui **consiste à utiliser un ensemble de techniques de manière à rendre impossible, en pratique, toute identification de la personne par quelque moyen que ce soit et de manière irréversible.**” CNIL*

Le moyen d’empêcher l’identification future.

La notion d’“effort raisonnable” (requérant 26) ⇒ identification

Activité : Réfléchir de quelle manière il est possible d'identifier des individus via une base de données.

Voici quelques informations sur des personnes publiques. Les reconnaissez-vous ?

- Un chanteur qui chante son énervement suite à la fermeture de son restaurant.
- Une ancienne enseignante de français pour des adolescent.e.s. et qui est la compagne d'un président.
- Un.e proviseur.e du lycée Diwan à Vannes.



Ce que nous conseillons aux équipes de recherche :

S'approprier les données :

Prendre connaissances de la méthodologie, de l'environnement des données...

Identifier les données potentiellement identifiantes et les données délicates et/ou sensibles.

Supprimer les données directement identifiantes :

Noms, prénoms... Tout ce qui est unique pour un individu unique.

Arbitrer pour les données indirectement identifiantes :

Entre la nécessité d'anonymiser et la pertinence scientifique.

Les pratiques que nous préférons éviter :

- Randomisation
- L'utilisation d'Alias

⇒ Eviter la transformation de l'information en une autre pour éviter les mésinterprétations

Le choix de la rédaction :

La généralisation :

- L'agrégation de données :
 - Prendre l'échelle du dessus, regrouper les modalités...
- Les Hyperonymes
 - *“C'est ainsi que ((PRENOM)) a rencontrée ((NOM DE L'ENTREPRISE))”*
- La combinaison de ces outils
 - *“Mais vous savez l'entreprise c'est ((ENTREPRISE DU BTP EN ESSONNE))”*

Réduire l'accès

- Comité déontologique
- Mettre un “contrat à la réutilisation”
- Mobiliser les outils de la plateforme, accès contrôlé...

Documenter l'anonymisation

- Expliciter les techniques utilisées
- Préciser les types d'informations modifiées
- L'exemple de l' [“Enquête sur l'enquête”](#) produit par le CDSP pour BeQuali :
 - Avec ceux du FN de Daniel Bizeul 1996-2003

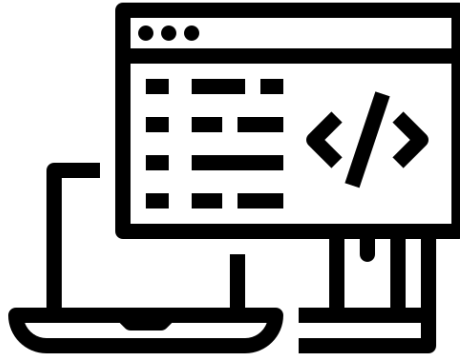
Les métadonnées et l'anonymisation

L'anonymisation sources de métadonnées.

Attention aux parodonnées des fichiers et métadonnées qui peuvent rompre la confidentialité.

Protégez vos données personnelles.

Des questions ?



Le catalogue de données du CDSP : Retours sur l'Expérience Utilisateur

Le catalogue du CDSP

Plus de 350 bases de données :

- **Résultats électoraux :**
 - Elections municipales, cantonales, départementales, régionales, législatives, présidentielles, européennes, référendums
 - Depuis 1958
- Enquêtes de recherche

Le catalogue du CDSP

Plus de 350 bases de données :

- Résultats électoraux
- **Enquêtes de recherche :**
 - Qualitatives (BeQuali) et quantitatives (Panel ELIPSS, CEVIPOF...)
 - Des approches disciplinaires variées : sociologie, sciences politiques, psychologie...
 - Des thématiques nombreuses : opinions politiques, pratiques culturelles, consommation, alimentation...

Le catalogue du CDSP

Plus de 350 bases de données :

- Résultats électoraux
- **Enquêtes de recherche :**
 - Quelques exemples :
 - Faire face au Covid-19
 - French electoral study 2017
 - Enquêtes auprès des députés, cadres de partis politiques...
 - Comportements alimentaires, connaissances financières, pratiques numériques, jeux de société, musées...

Conséquence : des documents variés

17 Octobre 1958

JOURNAL OFFICIEL DE LA RÉPUBLIQUE FRANÇAISE

9503

AVIS, COMMUNICATIONS ET INFORMATIONS

Ministère des finances et des affaires économiques.

Avis aux importateurs de matériels d'équipement, demi-produits, ouvrages en métaux et pièces détachées originaires et en provenance des États-Unis et du Canada (p. 9527).

Avis aux importateurs de produits originaires et en provenance de Suisse (p. 9527).

Avis aux importateurs (contingents globaux) (rectificatif) (p. 9528).

Avis aux porteurs de certificats d'investissements 5 p. 100 1953-1954 (p. 9528).

Ministère de la santé publique et de la population.

Avis de vacance du poste de directeur économe de l'hôpital-hospice de Pacy-sur-Eure (Eure) (p. 9528).

Avis de concours pour le recrutement d'un économe à l'hospice départemental de Neurey-les-la-Demie (Haute-Saône) (p. 9528).

BANQUE DE FRANCE

Avis relatif aux taux d'escompte et des avances sur titres de la Banque de France (p. 9528).

Situation de la Banque de France (p. 9529).

Annonces (p. 9530).

ORDONNANCES

Ordonnance n° 58-964 du 16 octobre 1958 relative à l'élection des députés des départements d'Algérie à l'Assemblée nationale.

EXPOSE DES MOTIFS

Les électeurs des départements d'Algérie sont appelés, à l'occasion de la prochaine consultation, à élire des représentants au même titre que la métropole. La présente ordonnance prévoit que les élections auront lieu au scrutin de liste majoritaire à un tour, sans

Art. 2. — Les circonscriptions et le nombre de députés que chacune d'elles est appelée à élire sont déterminés conformément au tableau ci-annexé.

Art. 3. — Chaque liste comprend un nombre de candidats égal au nombre de députés à élire.

La répartition des candidats de statut civil de droit commun et des candidats de statut civil local qui doivent figurer sur chaque liste est fixée dans chaque circonscription, conformément au tableau ci-annexé.

Art. 4. — Sont élus les candidats de la liste qui a obtenu le plus grand nombre de voix.

TITRE I^{er}

DES DÉCLARATIONS DE CANDIDATURE

Art. 5. — Les candidats d'une liste dans une circonscription électorale sont tenus de faire une déclaration revêtue de leur signature. Les déclarations de candidature doivent indiquer :

- 1° Le titre et le signe de la liste présentée;
- 2° Les noms, prénoms, date, lieu de naissance, statut civil, domicile et profession des candidats.

Plusieurs listes ne peuvent avoir dans les mêmes circonscriptions le même titre ou le même signe ni être rattachées au même parti ou à la même organisation.

Art. 6. — Cette déclaration doit également indiquer sous peine d'irrecevabilité les noms, prénoms, date, lieu de naissance, statut civil, domicile et profession des deux personnes, l'une de statut civil de droit commun, l'autre de statut civil local, appelées à remplacer les candidats élus en cas de vacance du siège; elle doit être accompagnée de l'acceptation écrite des personnes appelées à remplacer les candidats.

Nul ne peut être désigné comme remplaçant s'il ne remplit les conditions d'éligibilité exigées des candidats.

Nul ne peut figurer en tant que remplaçant sur plusieurs listes de candidatures ni être candidat dans une circonscription et figurer comme remplaçant d'un candidat.

Art. 7. — Les déclarations doivent être déposées en double exemplaire auprès des commissions de circonscription instituées au titre IV de la présente ordonnance au plus tôt trente jours et au plus tard vingt et un jours avant celui de l'ouverture du scrutin.

Il est donné au déposant un reçu provisoire de déclaration.

Conséquence : des documents variés

Variables List

Dataset contains 214 variable(s)

File fr_cdsp_ddi_elipss_201406atpa							
#	Name	Label	Series Name	Format	Valid	Invalid	Question
1	UID_atpa	Identifiant interviewé	continuous	numeric-3.0	810	0	Identifiant interviewé
2	atpa_ORD..	Ordre passage questionnaire [AFAM][ASCEN]	discrete	numeric-1.0	810	0	Variable calculée
3	atpa_AFAM1	Une famille : un homme et une femme mariés sans enfant	discrete	numeric-1.0	789	21	Un homme et une femme mariés qui vivent ensemble sans enfant(s). Forment-ils une famille ?
4	atpa_AFA..	Une famille : un homme et une femme mariés avec enfant(s)	discrete	numeric-1.0	284	526	Et s'ils ont des enfants ?
5	atpa_AFAM2	Une famille : un homme et une femme pacésés sans enfant	discrete	numeric-1.0	786	24	Un homme et une femme pacésés qui vivent ensemble sans enfant(s). Forment-ils une famille ?
6	atpa_AFA..	Une famille : un homme et une femme pacésés avec enfant(s)	discrete	numeric-1.0	315	495	Et s'ils ont des enfants ?
7	atpa_AFAM3	Une famille : un homme et une femme vivant en couple sans enfant	discrete	numeric-1.0	781	29	Un homme et une femme qui vivent ensemble en couple sans enfant(s). Forment-ils une famille ?
8	atpa_AFA..	Une famille : un homme et une femme vivant en couple avec enfant(s)	discrete	numeric-1.0	397	413	Et s'ils ont des enfants ?
9	atpa_AFAM4	Une famille : deux femmes mariées sans enfant	discrete	numeric-1.0	778	32	Deux femmes mariées l'une à l'autre qui vivent ensemble sans enfant(s). Forment-elles une famille ?
10	atpa_AFA..	Une famille : deux femmes mariées avec enfant(s)	discrete	numeric-1.0	420	390	Et si elles ont des enfants ?
11	atpa_AFAM5	Une famille : deux femmes pacésés sans enfant	discrete	numeric-1.0	771	39	Deux femmes pacésés l'une à l'autre qui vivent ensemble sans enfant(s). Forment-elles une famille ?
12	atpa_AFA..	Une famille : deux femmes pacésés avec enfant(s)	discrete	numeric-1.0	422	388	Et si elles ont des enfants ?

Conséquence : des documents variés



Regardez-vous des chaînes étrangères non francophones ?

Oui, tous les jours ou presque

Oui, de temps en temps

Oui, rarement

Non, jamais

[PAGE PRÉCÉDENTE](#) [PAGE SUIVANTE](#)

Une enquête “Expérience Utilisateur”

- Née de la nécessité d'évaluer nos services dans une perspective UX
- Lancée en avril 2021
- Une première vague terminée en juin...
- Et une deuxième vague qui sera lancée en octobre

L'enquête **Expérience Utilisateur**

Des objectifs multiples :

- **Évaluer la réutilisabilité des enquêtes**
 - Qualité des données
 - Richesses de la documentation
 - Facilité à prendre en main les enquêtes
- Comprendre quels outils sont plus appréciés
- Déterminer des besoins non-satisfaits

L'enquête **Expérience Utilisateur**

Des objectifs multiples :

- Évaluer la réutilisabilité des enquêtes
- **Comprendre quels outils sont plus appréciés**
 - Nesstar
 - Base de questions
 - Site BeQuali
 - Data.sciencespo
- Déterminer des besoins non-satisfaits

L'enquête Expérience Utilisateur

Des objectifs multiples :

- Évaluer la réutilisabilité des enquêtes
- Comprendre quels outils sont plus appréciés
- **Déterminer des besoins non-satisfaits**

L'enquête Expérience Utilisateur

Méthodologie :

- Contact de personnes ayant téléchargé une ou plusieurs enquêtes depuis 2020
- Entretiens semi-directifs
- Recodage des entretiens en variables d'intérêts

L'enquête **Expérience Utilisateur**

Résultats :

- **Satisfaction générale des services :**
 - Qualité des données et métadonnées
 - Outils (base de question, BeQuali)
 - Plateforme Quetelet Progedo Diffusion
- Attentes supplémentaires

L'enquête **Expérience Utilisateur**

Résultats :

- Satisfaction générale des services
- **Attentes supplémentaires :**
 - Sensibilisation du corps académique aux plateformes de données
 - Jeux de données simplifiés pour l'enseignement

La création de jeux de données pédagogiques

- Principe :
 - Utilisation de données en contexte d'enseignement
 - Nécessité de simplifier la base de données et la documentation
- Mise en pratique

La création de jeux de données pédagogiques

- Principe
- **Mise en pratique :**
 - Sélection de quatre enquêtes de thématiques différentes
 - Création de la base de données et de sa documentation
 - Anonymisation

La création de jeux de données pédagogiques

- Principe
- Mise en pratique :
 - **Sélection de quatre enquêtes de thématiques différentes :**
 - European Social Survey 9
 - Faire face au Covid-19 (vague 1)
 - Styles de vie et environnement
 - Fécondité, contraception et dysfonctions sexuelles
 - Création de la base de données et de sa documentation
 - Anonymisation

La création de jeux de données pédagogiques

- Principe
- Mise en pratique :
 - Sélection de quatre enquêtes de thématiques différentes
 - **Création de la base de données et de sa documentation :**
 - Sélection d'environ 50 variables
 - Création d'une documentation restreinte au format csv
 - Anonymisation

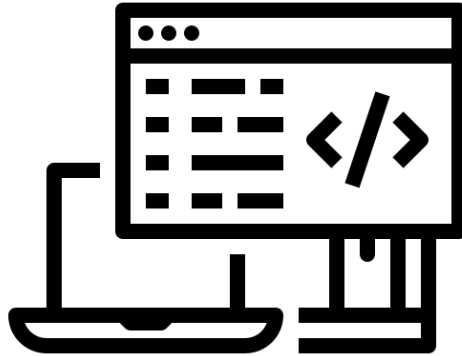
La création de jeux de données pédagogiques

- Principe
- Mise en pratique :
 - Sélection de quatre enquêtes de thématiques différentes
 - Création de la base de données et de sa documentation
 - **Anonymisation :**
 - Recodage de modalités
 - Suppression de données pour respecter le k-anonymat

La création de jeux de données pédagogiques

	A	B	C	D	E	F
1	Label	Code origine	Univers	Question	Modalités	Notes
2	covid_infection	coco1_q09		Avez-vous été atteint(e) par le virus Covid-19, que vous avez été testé(e) positif(ve) ou	1 : J'ai été testé(e) positif(ve) 2 : J'ai de fortes suspicions d'avoir été(e) atteint(e) 3 : Je ne pense pas avoir été(e) atteint(e) 9999 : N'a pas répondu	
3	covid_sante_post_infection	coco1_q10	Si la personne a été infectée par le covid-19 ou a de fortes suspicions (si Q09==1 ou Q09==2)	Comment vous portez-vous maintenant ?	1 : Je suis guéri(e) 2 : Je suis encore malade mais je vais plutôt bien 3 : Je suis malade 6666 : Non-concernée Valeurs numériques	
4	covid_infection_proches	coco1_q12	Si la personne a de la famille ou des connaissances atteintes par le Covid-19 (question externe à la t	Combien de ces personnes de votre famille ou de vos connaissances ont été testées posi	6666 : Non-concernée 9999 : N'a pas répondu Valeurs numériques	
5	covid_decès_proches	coco1_q14	Si la personne a de la famille ou des connaissances atteintes par le Covid-19 (question externe à la t	Parmi ces personnes, combien d'entre elles sont décédées aujourd'hui ?	6666 : Non-concernée 9999 : N'a pas répondu	
6	covid_experience_confinement	coco1_q62		Comment décririez-vous votre expérience du confinement ?	1 : C'est une expérience qui nous fait perdre beaucoup de temps 2 : C'est une expérience qui permet de prendre du recul et réfléchir à son mode de vie 3 : C'est une expérience qui permet de consacrer du temps à des choses essentielles comme la famille, les amis et les enfants 4 : C'est une expérience qui fatigue les nerfs 9999 : N'a pas répondu	

Des questions ?



Introduction aux paradonnées

Définition

Le terme de parodonnées a été forgé par Mick Couper (1998) dans le domaine des enquêtes.

Les parodonnées sont des données recueillies en parallèle d'un dispositif de collecte et qui en **décrivent le processus**, là où les métadonnées décrivent les données collectées.

Dans certains cas la distinction parodonnées / métadonnées s'avère difficile.

Rejoignez nous sur mentimeter :



Par lien : <https://www.menti.com/y6ue5en28j>

Par code : <https://www.menti.com/> puis utiliser le code 1369 3884

Sources de parodonnées

Les parodonnées sont de nature et de formats très divers, et peuvent représenter des volumes considérables..

Les sources de parodonnées dépendent du mode de collecte des données.

- les journaux d'appels des instituts de sondage
- les plateformes d'enquête (questionnaires auto-administrés)
- les applications mobiles utilisées comme instrument de collecte
 - les données recueillies par les capteurs des appareils mobiles (collecte passive de données)
 - données GPS
- [...]

Potentiel d'utilisation

Les parodonnées sont utilisées:

- pour piloter le dispositif de collecte
 - par exemple optimisation des déplacements d'enquêteurs
 - maîtrise des coûts (choix des séquence et des modes de rappel)
 - augmenter les taux de réponse
- pour améliorer la qualité de l'instrument de collecte
 - détection précoce des difficultés dans un questionnaire en ligne
- pour alléger la charge que la collecte fait peser sur les répondants
- pour nourrir des recherches méthodologiques

Exemple de paratonnées ELIPSS

Sources de parodonnées:

Journaux d'interactions

L'application enregistre certains événements :

- appels téléphoniques entrants et sortants
- campagnes de communication multi-mode (*)
- enregistrement des périodes d'indisponibilité (**)

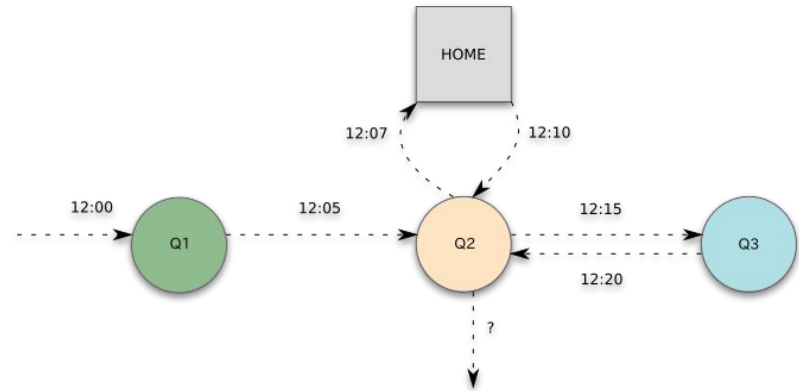


Sources de parodonnées:

Événements du questionnaire

Les traces d'activité (fichier de log), enregistrent :

- les dates et heures de passage
- certaines valeurs définies au moment de la programmation du questionnaire (**):
- certains événements liés spécifiques au mode d'administration (tablette) (***)



drafting response scenarii documentation, CDSP, 2018

Production des jeux de données

Difficultés prévisibles :

Sur la durée d'un projet, les méthodes et les procédures peuvent évoluer. Une étape d'harmonisation s'est révélée nécessaire pour traiter par exemple, dans le cas des tickets support ELIPSS ouverts par les gestionnaires de panel :

- les champs note en texte libre associés aux tickets
- l'évolution des labels de catégories de tickets

Production des jeux de données

Difficultés inattendues :

L'infrastructure informatique ELIPSS se divise en plusieurs éléments. Ce qui peut produire des ruptures inattendues.

- différences de Timezone,
- Valeurs manquantes (stabilité de la connexion 4G par exemple)

Freins à la publication des parodonnées

considération **économiques** :

- coût de préparation élevé
 - **compétences** nécessaires à la post-production
 - **temps** de traitement et **volume** des données

considération **éthiques et légales** :

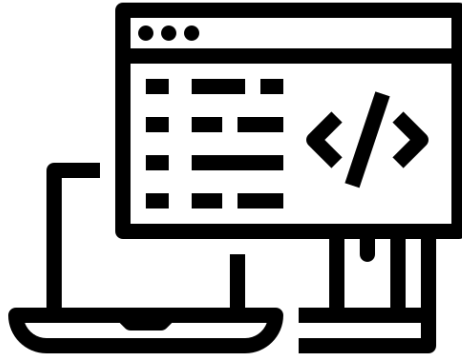
- problématique de ré-identification qui peut être difficile à résoudre
 - l'association des données de collecte principale et des parodonnées démultiplie les risques de ré-identification
- selon le RGPD, nécessité de recueillir **en amont** de la collecte, le **consentement éclairé** des participants à l'enquête

Freins à la publication des parodonnées

Un frein important à la publication des parodonnées, c'est l'absence de standard pour décrire de façon structurée les parodonnées .

De la même façon que les données primaires ont besoin de métadonnées, les **parodonnées ont besoin de métadonnées !**

Des questions ?



bibliographie

[Recherche sur la collecte des données à l'aide de paradonnées à Statistique Canada](#)

[Birth and Diffusion of the Concept of Paradata](#) Mick P. Couper Research Professor, Survey Research Center, University of Michigan

[Barkow, Ingo, Michaud, Geneviève, Perry, Anja, Schiller, David, & Thomas, Wendy. \(2020, November\). Paradata - from by-product to standard documentation. Presented at the 12th Annual European DDI User Conference \(EDDI20\), Zenodo. <http://doi.org/10.5281/zenodo.4299076>](#)

Bibliographie :

- Beaud, Stéphane et Weber, Florence, *Guide de l'enquête de terrain*, Paris, La Découverte, 2010. p.251.
- Coulmont, Baptiste. « Le petit peuple des sociologues. Anonymes et pseudonymes dans la sociologie française », *Genèses*, vol. 107, no. 2, 2017, pp. 153-175.
- Garcia, Guillaume. « Enquête sur l'enquête "Avec ceux du FN (1995-2003)" de Daniel Bizeul », juin 2020.
- Nguyen et Castelluccia, « Techniques d'anonymisation tabulaire : concepts et mise en œuvre ». 2020 : <https://arxiv.org/pdf/2001.02650.pdf>.
- [Guide de la bibliothèque de SciencesPo](#)
- [Guide UK Data Archives](#) / du [Finnish Social Science Data Archive](#)
- [Site internet de la CNIL](#)