

DATA CENSUS - INFOLAB

Guide de l'animateur

Sommaire

[1 - Le Data Census](#)

[Qu'est-ce qu'un Data Census ? A quoi ça sert ?](#)

[Pour qui ?](#)

[Comment ça marche ?](#)

[Quels participants ? combien de temps ?](#)

[Crédits et licence](#)

[2 - La méthodologie](#)

[Étape 1 : à partir d'un sujet, formuler des pistes d'usage : questions, besoins, défis \(30'\)](#)

[Étape 2 : première exploration de l'univers des données à partir de sources et de supports familiers \(30'\)](#)

[Étape 3 : la liste des données idéales : la wanted data list \(30'\)](#)

[Étape 4 : le Data Census proprement dit : les données véritablement existantes, celles qu'on imagine collecter et les autres \(60' ou +\)](#)

[Étape 5 : donner des perspectives : raconter un usage des données \(30'\)](#)

[Annexe 1 - kit du participant prêt à être imprimé](#)

1 - Le Data Census

Qu'est-ce qu'un Data Census ? A quoi ça sert ?

Un Data Census est une opération de **recensement des données¹ utiles, relatives à un sujet**.

Le Data Census permet de comprendre le paysage des données d'un sujet. A ce titre, il permet d'envisager des usages de ces données : répondre à des questions ou des besoins, démarrer un projet, etc.

Concrètement, le Data Census, à partir d'un sujet déterminé, permet de produire une liste de données la plus exhaustive possible :

- les données qui existent effectivement
- les données qu'il serait raisonnablement possible de produire
- les données idéales qui n'existe pas encore et sans certitude de pouvoir les produire

Cette liste de données permet de formuler une cartographie des données acquises et possibles pour pouvoir ensuite concevoir et développer des pistes d'usages.

Le Data Census n'est pas un objet figé et peut servir à des stades plus ou moins avancés de la médiation aux données.

Il peut servir très tôt pour explorer les données liées à un thème/sujet ouvert.

Partir d'un thème ou d'une question ouverte (par exemple "les données de santé") permet de se laisser surprendre par la donnée. Ici on ne sait pas tellement ce que l'on cherche, ce à quoi l'on veut répondre, mais on veut d'abord observer la richesse des données d'un système, d'un secteur et en tirer des questions et défis a posteriori.

On peut aussi partir d'un besoin mûr, par exemple "Nous voudrions un outil de dialogue avec l'acteur public sur les questions de vélo" ou encore d'un défi comme "Réduire la consommation énergétique d'un établissement scolaire". Le recensement de données va avoir pour objectif de repérer les données les plus pertinentes pour imaginer des pistes de réponses.

Pour qui ?

Le Data Census est une méthodologie "tous publics". Il n'y a pas besoin de connaissances pointues pour y participer. Il peut servir à des publics très divers : citoyens, professionnels dans le cadre de leur activité, étudiants, chercheurs, journalistes, habitants, acteurs publics, etc.

¹ Pour les spécialistes, le terme de "donnée" est ici à entendre au sens de "métadonnée".

Comment ça marche ?

Ce document explique étape par étape la réalisation d'un Data Census. Il ne prétend pas constituer une référence absolue et vous pouvez l'adapter à votre contexte, à vos besoins.

Ce document est destiné à l'animateur de l'opération. Après cette partie introductive, la partie "méthodologie" décrit pas à pas les 6 grandes étapes du Data Census.

Le "kit du participant", quant à lui, offre à l'animateur la possibilité, optionnelle, de partir d'un kit prêt à l'emploi pour ses participants. Il s'agit d'un document à part, non-inclus dans ce présent document.

Quels participants ? combien de temps ?

Cette proposition de méthodologie est dimensionnée pour une taille de groupe et un temps raisonnable : nous comptons environ 30 personnes maximum (la taille d'une classe) pour environ une demi-journée.

Par expérience, il est difficile de faire un Data Census efficace en moins de 3 heures de temps, même sur des sujets très resserrés : chaque étape demande une bonne dose de créativité et de réflexion, et donc de temps.

De même, un Data Census comptant moins de 3 personnes risque d'être peu productif et long. Le groupe gagne par ailleurs à être hétérogène, de manière à garantir une plus grande diversité créative :

- des usagers de services correspondant au sujet traité
- des participants qui connaissent le périmètre engagé par le thème (des professionnels, des habitants, des usagers, des collectivités, etc)
- de potentiels futurs réutilisateurs des données (start-ups, étudiants, ...)

Crédits et licence

Le Data Census est une méthodologie issue de la Campagne Infolab de la Fing. L'élaboration de ce document a débutée le 22 avril 2015 dans un groupe de travail regroupant : Marine Albarède, Sylvain Bouchard, Martin Cahen, Fabien Labarthe, Manon Molins et Pauline Reboul. Il a été ultérieurement retravaillé par Marine Albarède, Florence Bournonville, Renaud Francou, Manon Molins et Charles Nepote.

Ce document est mis à disposition selon les termes de la licence Creative Commons Attribution 3.0

France : <http://creativecommons.org/licenses/by/3.0/fr/>

Vous êtes libre de partager — reproduire, distribuer et communiquer — ce document, l'adapter et l'utiliser à des fins commerciales à condition de mentionner ses auteurs.

2 - La méthodologie

Étape 1 : à partir d'un sujet, formuler des pistes d'usage : questions, besoins, défis (30')

Le Data Census porte sur **un sujet partagé par tous les participants** -- par exemple, *ma santé* ou bien *la consommation d'énergie*.

Bien choisir son sujet ?

Un sujet large comme "la santé" peut produire des dizaines de pistes d'usages utilisant des centaines de données. Le risque d'un sujet trop large est de rendre le Data Census très long ou frustrant. Une bonne stratégie peut être de décomposer un sujet en 2 à 4 sous-sujets et de répéter le Data Census pour chaque sous sujet.

il est généralement plus intéressant de se restreindre à un sous-ensemble du sujet mais d'aller au bout du Data Census plutôt qu'un sujet trop large qui va laisser les participants au milieu du gué.

Le sujet est clairement posé, partagé avec l'ensemble des participants et idéalement affiché ou rappelé tout au long du Data Census.

Les participants imaginent et formulent alors des besoins, des questions, voire des défis relatifs au sujet. Ils sont exprimés de manière concise (pas plus d'une phrase).

Cette étape permet aux participants de se motiver en leur faisant imaginer et comprendre ce qu'ils vont pouvoir analyser voire résoudre grâce aux données. Au cours des étapes suivantes, ces besoins/questions/défis vont servir de support pour imaginer les données utiles à rechercher ou collecter.

Bien reformuler son sujet pour susciter des réponses créatives ?

Un sujet posé de façon un peu brute pourra donner des résultats médiocres. il peut-être utile de formuler le sujet sous la forme d'une question qui appelle et stimule des réponses créatives.

Par exemple :

- au lieu du sujet "ma santé", préférer une interrogation comme "Qu'est-ce qui ne se passe pas toujours idéalement dans votre parcours santé au quotidien, et qui pourrait être amélioré ?"
- au lieu du sujet "la transition énergétique", préférer une interrogation comme "Comment les individus peuvent-ils être partie prenante de la transition énergétique ?"

Livrable de l'étape : Une série de questions, défis, besoins, chacun matérialisé par une phrase courte.

Support(s) : Liste de questions/défis/besoins.

Temps imparti : 30 minutes ou environ 1/6 du temps total

Exemple de liste de questions/défis/besoins.

Sujet : *Qu'est-ce qui est important pour vous et qui ne se passe pas toujours bien et qui devrait être amélioré dans mon quotidien de santé et bien-être ?*

Prise de rendez-vous

- *être en contact avec mes médecins : par exemple pour obtenir facilement le code d'entrée de l'immeuble du cabinet médical*
- *Prendre des rendez-vous plus simplement (par e-mails, sms)*

Récupération de doc / résultats / Accès aux informations

- *Récupérer mes résultats directement, ne pas avoir à aller les chercher*
- *Obtenir copie de prescriptions, analyses, pour les conserver.*
- *Avoir mon certificat d'aptitude au sport sans avoir à refaire tous les ans un checkup complet qui ne se base même pas sur le checkup de l'année dernière*

Trouver un médecin / professionnel de la santé

- *Lors d'un déménagement par exemple ("je ne prends pas de nouveau client")*
- *Au meilleur rapport qualité/prix (dépassement honoraires and co.)*
- *Des rendez-vous via des plateformes de e-santé*

Trouver facilement toutes ses infos de santé / Gérer ma paperasse

- *Trouver / garder ses infos de santé numériquement*

Accéder à des informations / clés de compréhension

- *Accéder à des éléments qui permettent d'avoir une distance critique vis à vis des injonctions (retro-injonctionning)*
- *Sur mes droits et mes possibilités : je fais quoi si mon médecin s'est planté ?*
- *Sur mes remboursements (qu'est-ce que je peux me faire rembourser ?)*
- *Sur la santé d'un tier (un parent âgé qui ne peut plus vous dire ce dont il a besoin...)*
- *Sur mes allergies par ex. (contextualisation, etc.) Allergies croisées*

Suivi / Rappels / mise à jour

- *Envoyer des rappels, des mises à jour...*
- *Suivi actualisation carnet de santé, vaccins, etc.*
- *Le Pollen : la liste de mes allergies, comment les prévenir (prévoir date où le pollen va être fort par exemple) ?*
- *Avoir un système pour empêcher mon ordinateur de démarrer si j'ai pas fait ma gym (je la fais jamais)*

Étape 2 : première exploration de l'univers des données à partir de sources et de supports familiers (30')

L'objectif est de démarrer l'exploration de l'univers des données à travers des objets qui sont familiers à tous les participants -- spécialistes comme non-spécialistes du sujet. Il s'agit d'une première photo à gros grains ou bien un travail très pointilliste qui va servir de support aux étapes ultérieures : vérifier qu'on a pas oublié certaines pistes, certains territoires. **Cette étape n'a pas besoin d'être très rigoureuse et de parvenir à un résultat parfait.** Toutes les réponses données sont notées, même si elles sont un peu hors sujet, imprécises, voire même trop précises.

La question posée aux participants peut se formuler simplement : **où trouver des données relatives au sujet traité ?**

Pour aider les participants à y répondre, il est possible de prendre un angle encore plus concret : **quels sont les supports qui contiennent des données relatives au sujet : objets, documents, affichages, médias, page web, etc. ?** (Voir les exemples donnés ci-dessous.)

Cette question permet à des non-spécialistes de valoriser leurs savoirs informels, en démarrant ainsi sur un pied d'égalité avec les spécialistes. Exemple sur le thème de la santé : "je regarde les emballages des produits alimentaires, ils contiennent les ingrédients [des données !], me permettant de vérifier qu'il n'y en aucun auquel je suis allergique".

Cet exercice sur les supports peut suffire à cette étape. Si vous êtes dans des conditions idéales (pas trop nombreux et avec du temps devant soit) vous pouvez aussi aller un peu plus loin prendre d'autres angles d'exploration :

- Quels univers de données explore-t-on ?
- De quelles données parle-t-on ?
- Où peut-on aujourd'hui trouver ces données ?

Par exemple, sur le sujet "des individus partie prenante de la transition énergétique" :

- Données issues des compteurs et capteurs des foyers, bâtiments, collectivités, entreprises...
- Données de mobilité
- Données de consommation quotidienne, d'électroménager

Jusqu'où aller dans la cartographie des types de données ?

A la fin de cette étape, la cartographie doit être organisée : les types ou les supports de données doivent être suffisamment clairs pour que tous les participants s'y retrouvent.

L'animateur ne doit pas hésiter à pousser les participants à être précis.

Par exemple, en prenant le cas des données de santé : "sur mon ordonnance, je peux trouver la prescription, le nom de mon médecin, le nom du médicament, la date de ma consultation."

Livrable de l'étape : Une première cartographie "type de donnée" (**voir ci-dessous**), qui peut prendre la forme d'une liste ou d'un schéma.

Support(s) : Sur un tableau blanc ou projetée via vidéoprojecteur

Temps imparti : 30 minutes

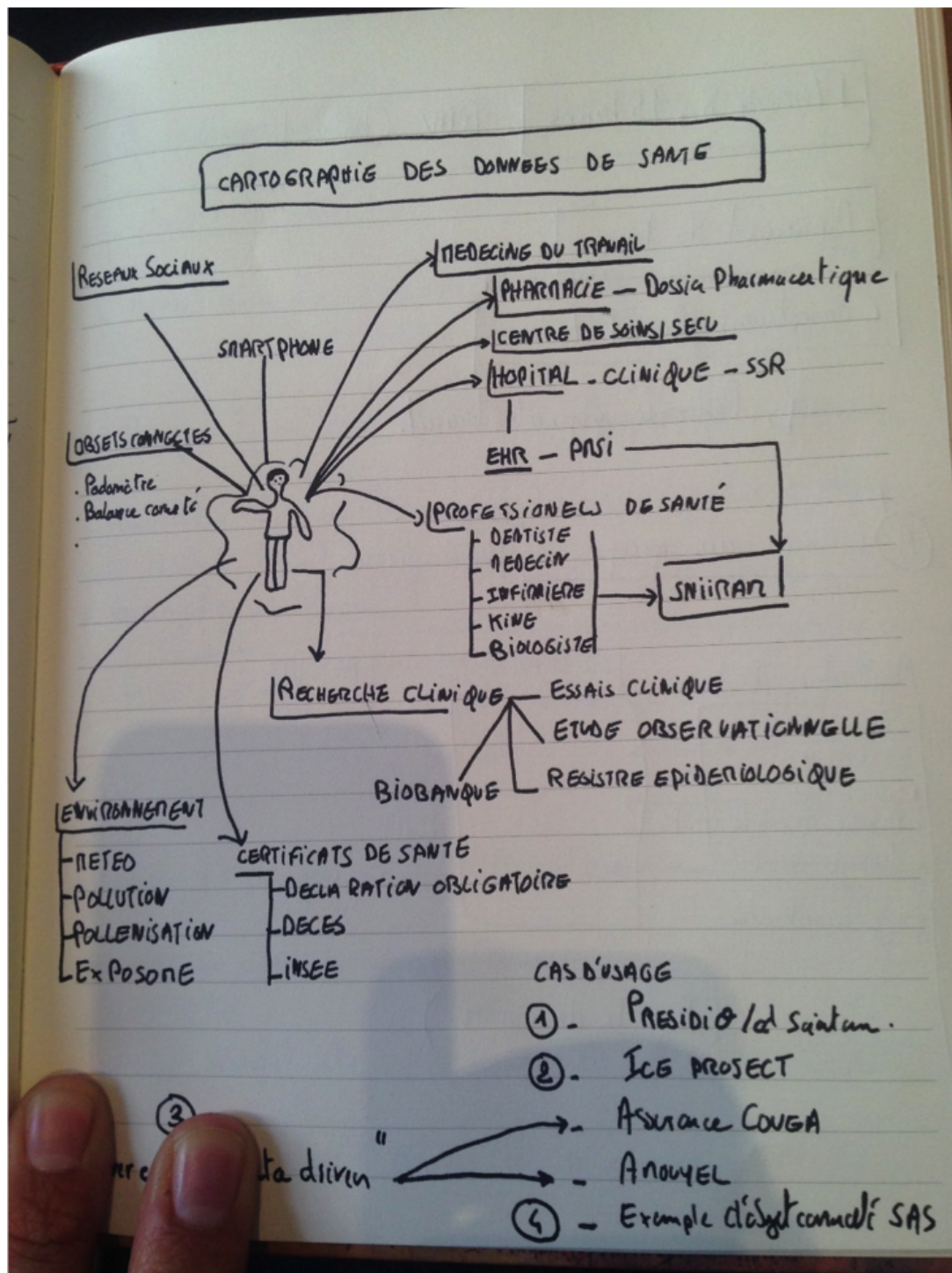
Exemple ce qui a été produit au cours d'une mini-session sur le thème de la santé personnelle (mai 2015) :

1. La CNAM a plein de choses...
2. Photos (ex du fils de Xxxx qui prend tout - ses documents - en photo)
3. Carnets de santé / le Dossier Médical Partagé (DMP)
4. Ordonnances (pdf) (correction des lunettes)
5. Ordonnances scannées par les pharmaciens
6. Résultats d'analyses
7. Radiographies, scanners...
8. Applications
9. Feuille de remboursement / historique de remboursement en ligne
10. Contrat mutuelle
11. Agenda médecin
12. Agenda perso
13. Codes barre sur boîte de médoc
14. Carte vitale
15. Historique des objets connectés (balance, piluliers intelligents, boîtier diabétique, toilette connectée, etc.)
16. Comptes bancaires
17. Autotests alcoolémie, drogues, grossesses...
18. Check up salles de sport, programmes de fitness
19. Invalidité (avec caractéristiques → à voir)
20. Téléphone
21. Lister ce qu'il y a dans le Healthkit (les 72 data
https://developer.apple.com/library/ios/documentation/HealthKit/Reference/HealthKit_Constants/index.html#//apple_ref/doc/uid/TP40014710-CH2-DontLinkElementID_1)
22. La composition des médicaments que j'ai ingéré plus ou moins récemment
23. Cartes des groupes sanguin / Dossiers des donneurs
24. Avis d'arrêt de travail, certificat médical (CERFA)
25. Dossier pharma
26. Compte Ameli
27. ADN
28. Données de généalogie
29. Carnet d'adresse => mes contacts santé (mon réseau)

Pour contextualiser tout ça :

1. acte de naissance (sexe et âge)
2. taux de pollution
3. pollens
4. la rubrique météo de mon journal quotidien
5. tickets de caisse (consommation alimentaire ; parapharmacie ; pharmacie)
6. menus des repas de la cantine scolaire/de l'entreprise.
7. Produits alimentaires / fiches produits OpenFoodFacts
8. Emballages cosmétiques que j'utilise (crème solaire, produit de beauté, etc.)
9. Smartphone (géolocalisation)
10. appareils de mesure du CO2 chez soi
11. données localisation antennes relais
12. données d'autres patients (patientslikeme)
13. Données liées au surbooking medical

Autre exemple d'une liste "type de données" (ici, la santé)



Étape 3 : la liste des données idéales : la wanted data list (30')

La "Wanted Data List" est la liste des données idéales qui seront utiles ou que l'on pense en rapport avec notre sujet.

C'est une manière de prendre pied avec le sujet. C'est un exercice simple au cours duquel les participants sont souvent très créatifs.

Comment créer cette liste ?

- Poser les bonnes questions :

"où sont les postes de consommation d'énergie dans mon collège?"

- ☐ trajets
- ☐ chauffage, éclairage
- ☐ les données du bâtiment (Quelle est l'isolation ? Le nombre de personnes dans le bâtiment ? Marche-t-il au gaz ou électricité ? A quels moments le bâtiment est utilisé?)
- ☐ position géographique

- Spécifier la donnée dans ces questions

A ce stade, il convient de décrire les jeux de données de la manière la plus simple, sans a priori ni référence à un jeu ou un fournisseur de données particulier : c'est la nature des données qui nous intéresse.

Il faudra cependant indiquer où on pense pouvoir la trouver, afin de faciliter l'étape suivante.

Par exemple : isolation => indice, nom d'un matériau, etc

Livrable de l'étape : un tableau simple (voir ci-dessous), qui référence toutes les données auxquelles on puisse penser, et les spécifier sur deux champs : le nom et où elle pourrait se trouver. Le photocopier, le distribuer aux participants pour que chacun réfléchisse de son côté, puis mise en commun.

Support(s) : Tableau blanc ou vidéoprojecteur pour la mise en commun

Temps imparti : 30'

Mes supports et actions

Supports :

Actions : _____

Exemple d'une Wanted Data List remplie :

Titre du projet : "Réduire la consommation énergétique de mon lycée"
Participants :, ...
Données idéales
Concentration de pollution dans la classe
Types polluant/pollutions : CO, benzène, parfums, désodorisants, bougies, cheminées, produits d'entretiens, ventilation, etc.
Habitants : fumeurs, âge, présence
Chauffage (type, consommation, bois ?)
Volume des pièces
Date de construction
Date des normes
Date des mises aux normes
Type de ventilation
Revêtements (murs, sol, peinture, mobilier (colle, placos (solvants), vernis)))
Matériaux
etc....

Étape 4 : le Data Census proprement dit : les données véritablement existantes, celles qu'on imagine collecter et les autres (60' ou +)

Temps 1 : Décrire les moyens possibles pour rechercher les données, ainsi que d'éventuels supports/documents où elles pourraient se trouver

Partir d'une énumération de moyens et chercher ensemble, selon le défi et le terrain, d'autres moyens.

- Open Data
- Coup de fil à un ami, au lieu de me déplacer, j'appelle pour trouver un interlocuteur qui me renseigne sur où trouver la donnée.
- Si elle n'existe pas : compter soi-même, mesurer soi-même (capteurs etc)
- Questionnaires
- Photos
- etc.

Livrable de l'étape : Utiliser la liste des moyens déjà disponible sur la méthodologie de recherche de données ([voir annexe](#)) et l'enrichir en fonction de votre sujet !

Support(s) : Tableau blanc ou vidéoprojecteur

Temps imparti : 15 minutes

Temps 2 : Rechercher les données

Les rassembler dans un tableau 4 à 5 champs: nom de la donnée, nom du détenteur, typologie de la donnée (date, texte, etc), courte description, et enfin la donnée.

Ce tableau sera en ligne, dans un document partagé pour se séparer le travail de recherche/collecte pendant l'atelier.

Livrable de l'étape : Tableau des données (voir ci-dessous), à remplir en ligne

Temps imparti : La recherche peut être longue.

Exemple (TODO: à rédiger) : thème retenu : ma santé

Nom de la donnée et/ou description sommaire	Détenteur/Source de la donnée (Où)	Mode de collecte possible
Liste des médecins de Marseille par leur prénom et nom	Ville de Marseille, liste papier du service santé	--
Liste des médecins de Marseille par leur prénom et nom	Pages Jaunes	--
Pour chaque médecin, accessibilité du cabinet en fauteuil roulant	-- [non trouvée]	Relevé manuel de terrain sur la base de la liste des médecins

Aller plus loin : visualiser l'ensemble du Data Census

Il est utile d'avoir une vision globale des données recensées. Le Data Census, un tableau un peu froid, peut être très frustrant et difficile à lire ! Il est donc intéressant de créer des représentations de ces données afin de créer du sens pour les participants.

Livrable de l'étape : Infographie, cartographie, dataviz etc.

Support(s) : Tableau blanc ou vidéoprojecteur pour le tableau qui rassemble les données et pour exposer le travail de dataviz (cartographie, infographie etc)

Étape 5 : donner des perspectives : raconter un usage des données (30')

Après le recensement, nous aurons des données qui permettront de proposer des pistes d'usage plus précises.

Il ne s'agit pas dans cette courte étape de décrire des scénarios d'usage, des services mais bien un ou deux usages possibles que l'on peut faire des données.

Par exemple : la donnée "hauteur du lycée" permet de voir si il est possible d'y placer des panneaux solaires.

Livrable de l'étape : Un usage des données

Support(s) : ???

Temps imparti : 30 minutes

Annexe 1 - kit du participant prêt à être imprimé

Le kit du participant fait l'objet d'un document à part, disponible à cette adresse.

https://docs.google.com/document/d/1GJFrCogheSdY2T2EdXZ1EMOYGIB6g_n9AgtfT2tRXrU/edit#heading=h.r993ia800w9