



PILOTE MESINFOS ENGAGER LE SELF DATA

DOCUMENTATION DU PILOTE - 1ere version : Mise en route du pilote

Juin 2017 - *Document intermédiaire*

Ce document vise à partager les conditions de réalisation du pilote MesInfos, les difficultés rencontrées, les points d'attention, etc.

Cette première partie documente l'amorçage du projet, essentiellement sur l'année 2016 et le début de l'année 2017. Il sera mis à jour au cours de l'année 2017.

Table des matières

DOCUMENTATION DU PILOTE - 1ere version : Mise en route du pilote	1
I. Introduction - Engager le Self Data - le choix du format du pilote	2
II. Constitution du tour de table	4
III. Restitution des données	6
IV. Plateforme Self Data - de Cozy à plusieurs plateformes ?	8
V. Données et services - réutilisation	12
VI. Recrutement du panel	14
VII. Observation et recherche	15

I. Introduction - Engager le Self Data - le choix du format du pilote

Pourquoi nous avons souhaité faire cela ? Et pourquoi en faire un projet collectif ?

La Fing a lancé le projet MesInfos en 2012, afin d'explorer la piste du retour des données personnelles aux individus, ainsi que ce que nous avons plus globalement nommé le Self Data : permettre aux individus de collecter, visualiser, comprendre, faire usage de leurs données personnelles, sous leur contrôle et à leurs propres fins. Les enjeux d'une telle piste sont multiples : en finir avec une asymétrie informationnelle entre les organisations et leurs clients, qui pèse sur la relation de confiance ; redonner des capacités d'action aux individus afin de construire une économie plus durable, en allant dans le sens des évolutions réglementaires (droit à la portabilité du GDPR, par exemple)...

Au cours des années, nous avons travaillé à identifier les enjeux, les risques et opportunités d'une telle piste ; nous avons mené une première expérimentation, en 2013-2014, qui rassemblait des organisations publiques et privées. A cette occasion, durant 8 mois, plus de 200 individus avaient accès à certaines de leurs données personnelles (banques, télécommunication, tickets de caisse...), qu'ils pouvaient utiliser comme ils le souhaitaient grâce à des services tiers.

Depuis, le contexte social et réglementaire a évolué : la défiance des individus quant aux usages faits de leurs données par des tiers s'est accrue (se traduisant notamment par l'augmentation de l'usage des adblockers), des services et "PIMS" émergent, le GDPR a instauré le droit à la portabilité que les organisations devront mettre en oeuvre d'ici mai 2018...

Le format du pilote

Dans ce contexte, le projet MesInfos se fixait comme objectif en 2016 et 2017 de "Faire franchir un cap" au Self Data. Et pour ce faire, il nous fallait davantage connaître et valoriser son potentiel, et donc obtenir des retours d'usage, instruire les défis techniques et juridiques. Nous avons donc décidé de lancer un pilote qui visait notamment :

- À appréhender et mesurer la valeur du Self Data pour les individus (perception de la plateforme de cloud personnel, des données personnelles, usages ou non-usages de services permettant aux individus d'utiliser leurs données personnelles, etc.)
- À accompagner la mise en oeuvre (technique, juridique, opérationnelle...) de la restitution des données personnelles aux individus par les organisations, et élaborant des process et canaux de mise à disposition des données, en traitant les questions juridiques ou techniques liées à ce sujet.

A noter, le droit à la portabilité (pas encore confirmé lors du lancement du pilote) va dans le même sens que le pilote sur ce dernier point ; en cela, MesInfos peut être vu comme un terrain / bac à sable pour commencer à mettre à disposition les données, étudier des formats et process de transmission des données qui pourraient être industrialisés, explorer les usages et services qui peuvent voir le jour...

Le pilote MesInfos a été conçu dès le départ comme un projet collectif (associant les partenaires aux grandes décisions et au pilotage du projet) ; le développement du Self Data repose nécessairement sur un écosystème, comprenant détenteurs de données, plateformes personnelles de données, services tiers, individus... nous sommes par ailleurs convaincus que la valeur du Self Data réside dans le croisement des données personnelles issues de divers secteurs par les individus, et donc dans le décroisement. Ces objectifs et questions ne nous semblent donc pouvoir être adressés que collectivement.

Quel format du pilote ?

Le pilote MesInfos prend la forme suivante :

- Des organisations (privées et publiques) sont rassemblées pour réfléchir à la mise en oeuvre du Self Data ; certaines sont prêtes à restituer des données, d'autres ont besoin de réfléchir à la mise en oeuvre d'une telle restitution, de se préparer. Nous avons ainsi assumé 3 niveaux de participation : "Explorateur", "Engagé" et "Pionnier".
- Des plateformes Self Data, permettant aux individus d'accéder à leurs données et de les réutiliser grâce à des services tiers : la pluralité des plateformes est notre ambition, essentielle pour faire émerger un marché qui plus tard, laissera le choix aux individus ; mais pour commencer, nous avons dû éviter la complexité inhérente à la multiplication des architectures des plateformes, en s'appuyant sur une plateforme en particulier, CozyCloud.
- Des services, proposés par des "réutilisateurs" divers : innovateurs, startups, étudiants, développeurs, partenaires détenteurs (qui peuvent jouer un double rôle).
- Un territoire living lab, pour mobiliser testeurs et réutilisateurs dans un même lieu et , qui est aussi la première collectivité à réfléchir activement au Self Data pour ses citoyens/usagers.
- Une équipe de chercheurs (sociologues, marketing, ergonomes...), qui observera et analysera différentes composantes du pilote : acteurs du Self Data (motivations, visions), individus (appropriation des données, de la plateforme, des services, usages, perception, etc.)
- Et enfin, évidemment, les individus, qui recevront leurs données ; nous avons prévu deux phases, une première avec des beta testeurs (300), une seconde avec une mobilisation plus grande (3000).

Comment animer un tel projet collectif ?

La Fing anime la réflexion globale et pilote le programme... tout en réfléchissant à une autonomisation progressive ; en effet, celui-ci n'a pas vocation à être porté par l'association sur la durée, le Self Data devant peu à peu être approprié par tous.

En 2016-2017, le pilote est animé selon une alternance de temps de partage et de construction collective (comités de pilotage, ateliers interpartenaires...) et de temps d'échange et de travail bilatéraux, par exemple sur la question des données. Pour cela nous avons défini un certain nombre de chantiers, interdépendants mais menés selon des processus bien distincts avec les différents partenaires du projet :

- Chantier données : engager la restitution des données à ses clients ou usagers
- Chantier clients : recrutement (propre au pilote) et expérience client/usagers
- Chantier réutilisation / services : faire émerger des concepts et des prototypes de services qui pourront être testés par les individus participant au pilote.
- Chantier observation/recherche : animation des testeurs, observation et analyse des retours d'usage et d'expérience.
- Chantier plateforme : amélioration du parcours utilisateur sur la plateforme, réflexions autour de la participation d'autres plateformes

En 2016, l'équipe MesInfos comptait cinq personnes : trois sur la coordination générale et le suivi de certains chantiers "partagés" (clients, plateforme, réutilisation), un coordinateur technique du projet, une personne en charge de la coordination du chantier "réutilisation", une personne en charge de la coordination du chantier "observation/recherche".

II. Constitution du tour de table

Le tour de table du pilote comptait à partir de 2016 des acteurs variés : outre la Fing, il comprend Cozy Cloud (plateforme) et les partenaires privés et publics : EDF, Enedis, GRDF, Engie (énergie), la MAIF et la MGEN (assurance), Orange (télécommunications), le Crédit Coopératif, RMW (Banque), Qwant, le Grand Lyon (collectivité), NTT Data... Plusieurs niveaux d'engagement ont été prévus ; en 2016, 3 de ces partenaires restituaient effectivement des données, 6 en 2017, d'autres partenaires étant davantage dans une posture d'explorateurs du sujet.

D'autres partenaires sont associés au pilotage en vertu de leur expertise : CNIL, Tuba, Cap Digital (acteurs de l'innovation et pôles de compétitivité...), etc.

Comment constituer le tour de table pour un tel projet ?

La constitution d'un tel tour de table ne se fait pas du jour au lendemain. Dans le cadre de MesInfos, nous nous sommes appuyés sur une implication des partenaires dans la durée d'une part (certains partenaires suivant MesInfos depuis plusieurs années) mais aussi sur de nouveaux partenaires ayant rejoint le projet en 2016. Si les choses sont très différentes selon les organisations, voici quelques étapes "inévitables" :

- Identifier la bonne entrée (la personne dont les thématiques de travail ou le métier font écho au Self Data, qui soit en mesure de faire monter le sujet, d'embarquer les bonnes personnes en interne)
- L'outiller pour convaincre en interne
- Prévoir plusieurs rencontres et rendez-vous, avec différents interlocuteurs
- Organiser un vrai temps d'accueil sur le projet / de remise à niveau lorsqu'un nouveau partenaire rejoint le projet

Quelles raisons de s'engager dans le self data, quels sont les principaux arguments ?

Les bénéfices attendus sont clairs du côté des individus ; récupérer leurs données personnelles peut-être l'occasion de sortir d'une situation de non-valeur, dans laquelle les

organisations collectent et traitent énormément de données sur leurs clients, alors que les bénéfices de cette collecte sont peu évidents pour ces derniers. Les bénéfices du Self Data pour les individus peuvent être pratiques (gagner du temps, se faciliter la vie, accéder rapidement à la bonne information ou à la bonne personne en fonction de ma situation...), économiques (comparer les tarifs, trouver l'offre qui me correspond, etc.).

Pour **les organisations détentrices de données**, s'engager dans le Self Data et plus particulièrement dans un projet pilote sur ce sujet peut être une opportunité de :

- Renforcer ou recréer le lien de confiance avec leurs clients ou usagers, en sortant d'un usage unilatéral des données, en élaborant de nouveaux canaux d'échange.
- Améliorer la qualité des données ; des individus auraient tout intérêt à tenir leurs données à jour s'ils en étaient les utilisateurs.
- Proposer de nouveaux services à leurs clients ou prospects, qui permettraient à l'individu de croiser diverses données sous son contrôle.
- Se préparer à la mise en oeuvre du droit à la portabilité instauré par le GDPR, en testant des canaux de mise à disposition, en apprenant sur les enjeux et défis juridiques, techniques, opérationnels, etc.

Quel rôle de la CNIL (l'autorité française de contrôle en matière de protection des données personnelles) dans ce projet ?

La présence de la CNIL dans le pilote MesInfos est indispensable à plusieurs titres :

- Elle apporte son expertise du sujet des données personnelles, en participant à la majorité des réunions transverses et comités de pilotage. Le service des Etudes et de la Prospective est le principal interlocuteur transverse sur le projet, qui s'inscrit dans les deux objectifs que vise la CNIL dans le cadre de son activité d'études, d'innovation et de prospective : *"détecter en amont de nouveaux usages et tendances émergentes et explorer des sujets prospectifs touchant aux libertés individuelles et publiques, aux données personnelles et à la vie privée dans l'univers numérique"*
- Elle accompagne les réflexions autour de la mise à disposition des données et sa mise en oeuvre opérationnelle. Plusieurs réunions ont ainsi été organisées en 2016 entre la CNIL et les partenaires du projet MesInfos, afin de construire un cadre juridique solide partagé. Ces réunions sont également indispensables pour rassurer les partenaires, pour qui il s'agit ici d'expérimenter quelque chose de neuf, sur le sujet de la chaîne de responsabilités, du fonctionnement des plateformes de cloud personnel, du lien avec les évolutions réglementaires (ex. GDPR)

III. Restitution des données

Comment s'y est-on pris ? Un interlocuteur au sein de l'équipe du pilote MesInfos, a impulsé le chantier données chez les partenaires, et les a accompagnés dans les différentes étapes :

1. Formuler une liste des données : établir une liste de ce qu'il y a dans vos systèmes d'information en termes de données, que vous envisagez de transmettre aux individus qu'elles concernent (idéalement en mobilisant des collaborateurs de la DSI - voire de la gouvernance de la donnée - des métiers, habitués à travailler avec ces données au quotidien, et de la relation client / marketing)

2. Initier une procédure d'approbation de l'initiative, du principe et des modalités de transmission des données, et de la liste des données par votre organisation.

Tous les partenaires qui transmettent des données en 2016 dans le pilote MesInfos ont dû passer par là. Ces procédures internes étaient dans chaque cas différentes, mais en général assez longues, et demandent de porter le message en interne dans différents services. Ici encore, associer des acteurs extérieurs peut être opportun.

3. Formaliser la liste des données, et commencer la documentation. Cela permettra notamment de l'ajouter dans un outil indispensable au pilote : <http://mesinfos.fing.org/cartographies/datapilote/>. Cet outil est un support d'échange clé avec les acteurs du projet, et permet de faciliter l'appropriation des données par des ré-utilisateurs (cela mobilisera la DSI et les métiers pour élaborer la documentation précise).

4. Préparer un système de transmission : nous penchons vers des éléments automatiques, mais pour l'instant cela reste des solutions assez légères, dimensionnées pour les quelques milliers de testeurs du pilote.

Ici, les questions à traiter (qui mobiliseront principalement la DSI) :

- identification : comment relier l'individu qui demande ses données, à son identifiant client dans le SI (et ainsi à ses données) ?
- authentification : comment s'assurer que la personne qui demande ses données est bien celle qu'elle prétend être ?
- transmission : comment transmettre de manière sécurisée les données ?

En parallèle, un travail de fond a été mené avec la CNIL, pour décrire un cadre juridique précis quant à la transmission des données ; celui-ci s'est appuyé sur des rencontres entre des juristes des organisations partenaires et des juristes de la CNIL, et sur un travail de consolidation et de centralisation des questions adressées par les détenteurs de données à la CNIL. Cela a permis de formuler les questions juridiques cruciales pour la mise en oeuvre d'un tel projet et d'obtenir des réponses de la part de la CNIL. Ces éléments ne sont pas à négliger et peuvent ensuite devenir un outil puissant pour défendre le projet en interne, chez les différentes organisations.

Quelles sont les données concernées ? Les organisations doivent-elles restituer toutes les données qu'elles ont sur leurs clients ou usagers ?

Les organisations ne doivent pas restituer dans le pilote MesInfos toutes les données qu'elles ont dans leur système d'information. Les données concernées sont de trois types :

- Les données administratives et sur la relation client (CRM) : fiche client, données de segmentation, facturation, contrat, ... On les retrouve dans toutes les organisations, souvent avec quelques spécificités. Elles amènent des cas d'usages assez administratifs. Par exemple, le cas d'usage de mise à jour automatique de ces données, de l'individu vers l'organisation.
- Des données transactionnelles souvent très liées et spécifiques au métier de l'organisation. Ce sont les relevés de consommation d'énergie (énergéticiens), les relevés de comptes, les sinistres (assurances), le journal d'appel (télécom),... Elles offrent souvent un regard et un point de vue objectif à l'individu sur des actions dont il n'a pas forcément pleinement conscience. Ce sont ces données qui ouvrent les potentiels de cas d'utilisation, et la valeur d'usage des services du Self Data.
- Des traces et des communication et points de contacts entre l'organisation et son client, tels que les horodatages de connexion sur l'espace client Web, les dossiers de suivi du service client...

Les données issues de traitements effectuées par les organisations (par exemple scoring, segmentation) ne sont généralement pas restituées (même si rien n'interdit aux organisations qui le souhaiteraient de restituer ces données). Parmi les données évoquées plus haut, le choix des données restituées par chaque partenaire dans le cadre du pilote dépendra de plusieurs facteurs : valeur potentielle pour l'individu, degré de complexité des canaux de transmission à mettre en place, obligations réglementaires liées à certains type de données...

Quelles sont les modalités de transmission possibles ?

Plusieurs solutions de transmission sont possibles dans le cadre du pilote. Des modalités automatiques (API...) sont privilégiées. En 2016, certains partenaires ont proposé d'utiliser des API, prévues pour un autre usage initialement (application mobile, ...). Pour d'autres, une API d'identification et d'authentification existante a pu être utilisée, complétée par une API, mise en place pour l'occasion, où les cloud personnels de chaque individu viennent récupérer les données qui les concernent régulièrement (pull). À ce jour, ces APIs sont dimensionnées pour les quelques milliers de testeurs du pilote MesInfos.

Qui est responsable lorsque les données sont mises à disposition des individus ?

La responsabilité des détenteurs de données (responsables de traitement initiaux) s'arrête dès lors que les données sont dans le cloud personnel des individus. Une fois que ces données sont transmises, le responsable de traitement initial n'est plus responsable de ce qui est fait de cette copie des données par les individus. L'opération de la transmission des données, notamment concernant sa sécurisation, est par contre de la responsabilité de l'organisation.

Les solutions mises en oeuvre pourraient-elles être celles de la portabilité demain ?

Les solutions mises en oeuvre dans le pilote ne sont probablement pas exactement celles de la portabilité, pour deux raisons :

- Le périmètre des données restituées dans le cadre du pilote MesInfos ne sera probablement pas identique à celui du droit à la portabilité, même s'il s'en rapproche (on retrouve dans les deux cas l'idée que les données restituées sont surtout des données "fournies"/"provided").
- Les modalités mises en oeuvre dans le cadre du pilote sont pour la plupart automatiques, mais ne sont pas dimensionnées pour accueillir des demandes de transfert de plusieurs dizaines, centaines de milliers de personnes.

Néanmoins, pour mettre en oeuvre le droit à la portabilité, les organisations devront faire demain ce qu'elles font déjà dans le cadre du pilote MesInfos : identifier les données, mettre en place des modalités de transmission automatisées en s'interrogeant sur la faisabilité technique et opérationnelle et le cadre juridique, documenter ces données, etc. C'est en cela que le pilote peut être vu comme un bac à sable de la portabilité, et une manière de faire de la réponse à ce droit une source d'innovation et de nouveaux services

Cette documentation du chantier "Données" s'enrichira largement au cours de second semestre 2017, notamment via un groupe de travail portant sur ce chantier dans l'application de la portabilité.

IV. Plateforme Self Data - de Cozy à plusieurs plateformes ?

Comment s'y est-on pris ? Quel choix fait ?

Le récit du Self Data a besoin de plateformes personnelles de données pour exister : des individus, outillés par des plateformes personnelles de données (cloud personnels) et par des services tiers, peuvent récupérer, stocker et utiliser leurs données personnelles selon ce qui fait sens pour eux.

Le pilote MesInfos avait donc besoin de s'appuyer sur une plateforme, pour expérimenter le Self Data. Il a ainsi débuté en 2016 en s'appuyant sur Cozy, plateforme de cloud personnel porté par la startup française CozyCloud (cela avait déjà été le cas lors d'une première expérimentation 2013-2014). Pourquoi cette plateforme plutôt qu'une autre ? Parce que Cozy était un des rares acteurs à ce niveau de maturité sur le marché français fin 2015-début 2016, présentant plusieurs caractéristiques (décentralisation, un cloud par individu, plateforme de services...) qui nous semblaient permettre de tester à la fois la valeur du retour des données aux individus, l'appropriation d'un tel espace, la valeur d'usage et les usages possibles de ces données... De plus, CozyCloud était déjà bien connu des partenaires de MesInfos, et avait des relations constructives avec plusieurs d'entre-eux ; ainsi, la Maif a-t-elle investi dans CozyCloud - indépendamment du pilote MesInfos. Dans le

cadre du pilote, c'est elle qui recrute les testeurs parmi ses clients et leur propose d'ouvrir leur propre espace de cloud personnel.

A moyen terme, le marché des plateformes Self Data doit être divers, afin de proposer plusieurs choix aux individus et leur permettre de changer de plateforme s'ils le souhaitent. Néanmoins, c'est pour éviter une complexité trop grande (diversité des architectures entre plateformes notamment) que nous avons choisi de nous appuyer sur une seule plateforme pour commencer, avant d'ouvrir à d'autres plateformes.

Ouvrir MesInfos à d'autres plateformes est compliqué ; une partie des API des partenaires pourrait être accessibles pour d'autres plateformes courant 2017, mais comme nous l'avons évoqué, cela prendra du temps avant que ces pratiques ne se généralisent.

- ***L'équipe MesInfos a produit au printemps 2017 [une note](#) afin de faire en premier pas dans cette direction.***

Cozy, plateforme du pilote MesInfos

Si Cozy est déjà un produit sur le marché, plusieurs étapes ont été nécessaires pour qu'il devienne une plateforme du pilote à proprement parler :

- Une validation auprès des partenaires, afin que ces derniers acceptent d'ouvrir les API existantes ou nouvellement créées aux connecteurs Cozy. Il s'agit surtout de s'assurer du respect de la confidentialité, de la sécurité et de la qualité de service de la plateforme.
- Une étape de création des connecteurs Cozy, permettant aux individus d'aller "requêter" leurs données
- Un travail sur le parcours usager de la plateforme Cozy, afin d'accompagner le passage de la communauté - déjà existante - des premiers beta-testeurs Cozy à des utilisateurs parfois moins à l'aise avec les technologies (qui allaient être ceux du pilote MesInfos). Cela nécessite d'être didactique, de penser un parcours simple - sans trop simplifier pour ne pas masquer les enjeux de la maîtrise de ses données - afin que les utilisateurs comprennent ce que signifie le cloud personnel et les possibilités d'utiliser leurs données.

Ce travail a été piloté par le coordinateur technique du pilote MesInfos pour tous les aspects techniques (accompagnement du développement des connecteurs, etc.), en lien avec d'autres personnes de l'équipe pour les autres sujets.

- **Les principaux points d'attention / de vigilance :**
 - La plateforme est un partenaire pivot du projet. Il est primordial que **la vision de la plateforme et la vision du projet concordent**, afin que la dynamique soit vertueuse, et que des synergies soient possibles.
 - Penser multiplateforme à ce stade de maturité du marché du Self Data est nécessaire, mais loin d'être évident. **Pour qu'une solution soit pleinement**



une plateforme Self Data, il nous semble nécessaire qu'elle remplisse au moins deux critères :

- Permettre à leurs utilisateurs d'accéder à toutes les données personnelles mises à disposition par les partenaires du pilote.
- Permettre, faciliter et encourager l'usage de services tiers, ré-utilisateurs des données personnelles des individus. Leurs fonctionnalités doivent donc aller au-delà des fonctionnalités classiques de stockage de documents ou de données des coffres-forts numériques.

- Mais pour expérimenter un marché du Self Data multiplateforme, il faudra lever les craintes des organisations détentrices du tour de table, afin qu'elles acceptent d'ouvrir leurs API non pas uniquement pour la première plateforme de test, mais pour plusieurs plateformes qui rempliraient des critères de sécurité et respecteraient les principes du Self Data. Cela peut passer par le fait de faciliter la rencontre entre ces plateformes et les détenteurs de données.

Quelles responsabilités de la plateforme de cloud personnel ?

La plateforme de cloud personnel prend la responsabilité d'héberger les données de l'individu : d'en assurer un hébergement fiable et sûr. Elle prend aussi l'engagement de lui permettre d'en faire usage, c'est à dire de les utiliser à travers de multiples services offerts par des tiers. Sa responsabilité est de donner les outils adéquats à l'individu pour qu'il ait conscience de l'utilisation de ses données et qu'il garde le contrôle sur ses données (en structurant l'accord de consentement avec les services tiers, en donnant des outils de contrôle des services, ...)

Qui est responsable lorsque les données sont entre les mains des individus ?

L'organisation qui a transmis les données n'est plus responsable : la responsabilité de traitement du détenteur s'arrête au moment où il n'a plus la maîtrise des données (en termes de finalité d'usage et de moyens de traitement) dans la mesure où il transmet ces dernières à la personne concernée.

Si les individus sont responsables de ce qu'ils font de leurs données, la loi française ne considère pas que l'individu puisse être "responsable du traitement" soumis à la loi s'il fait un usage personnel et domestique de ses données ; si l'individu a recours à des services tiers qui utilisent ses données, ce sont les fournisseurs de ces derniers qui sont responsables des traitements effectués sur les données.

Quelles conditions pour recevoir les données des partenaires ?

Bien que n'étant pas responsable du sort des données dès qu'elles sont dans le cloud personnel des individus (sous leur contrôle donc), les détenteurs restent attentifs quant au risque d'image : ces organisations ne peuvent pas se permettre de conseiller des services qui ne seraient pas de bonne qualité. Les points de validations portent ainsi sur l'hébergement (localisation, qualité, sécurité), les procédures en cas d'incidents, les traitements de données effectués par la plateforme, l'accompagnement des individus dans la gestion de leur données personnelles, ...

Quelle responsabilité pour l'application qui traite les données dans le Cozy (dont le code tourne dans Cozy) ?

Les fournisseurs des applications sont globalement responsables des traitements opérés par leurs applications sur les données des individus. Il y a une différence entre les applications qui tournent en local, sur le serveur, et les applications qui font sortir les données. Lors de l'installation des applications, l'utilisateur est informé de l'usage fait de ses données, notamment si elles sont traitées en local ou sortent de leur Cozy.

Dans quelle mesure les détenteurs de données sont tenus de s'assurer de la légitimité des plateformes sur lesquelles ils déversent les données de leurs clients ?

A priori, il n'y a pas de raison pour les détenteurs de vouloir contrôler les espaces sur lesquels les individus stockent leurs données - qui dépendent du choix de ces derniers. Néanmoins, ils doivent pouvoir s'assurer de la sécurité du transfert des données jusqu'à la plateforme.

Par ailleurs, quand le marché des plateformes sera plus important, il sera difficile pour les détenteurs d'auditer chaque plateforme. Mais dans le cadre du pilote MesInfos, un ensemble de précautions a dû être pris ; ces plateformes et la restitution de données étant quelque chose de neuf, Cozy a dû être présenté aux Commissions de sécurité des détenteurs de données restituant les données.

Quelles conditions pour qu'une plateforme puisse participer au pilote MesInfos après mi-2017?

Si le projet MesInfos est relativement agnostique sur l'architecture et le stockage des données (sur un terminal, un serveur distant, P2P...) permis par la plateforme, les plateformes qui participeront au pilote doivent remplir plusieurs critères :

- Respecter la charte Self Data, signée par toutes les parties prenantes du projet ; en particulier, redonner la maîtrise aux individus (en leur permettant de gérer et faire usage de leurs données personnelles depuis un point unique, sous leur propre contrôle).
- Permettre à leurs utilisateurs d'accéder à toutes les données personnelles mises à disposition par les partenaires du pilote.
- Permettre, faciliter et encourager l'usage de services tiers, ré-utilisateurs des données personnelles des individus.

V. Données et services - réutilisation

Comment s'y est-on pris ?

Le chantier de la réutilisation des données est un des plus ambitieux et complexes. Pour que le récit du Self Data fonctionne, il est nécessaire de faire émerger rapidement des services. C'est aussi la condition *sine qua non* pour tester différents cas d'usages avec les individus, un des objectifs du pilote. L'expérimentation MesInfos menée en 2013-2014 avait confirmé que le contrôle des données n'était pour la majorité des gens pas une promesse suffisante pour qu'ils se préoccupent de leurs données, et qu'il y avait un besoin de services, même basiques, pour que les individus puissent en faire quelque chose.

Ces services dont nous parlons peuvent venir de plusieurs acteurs : développeurs indépendants ou startups, étudiants, détenteurs de données eux-mêmes, communauté de beta-testeurs Cozy (l'ADN de Cozy repose aussi sur sa communauté d'early adopters qui peuvent également être des développeurs). Nous avons souhaité travailler sur ces différents axes et avec ces acteurs.

MesInfos a ainsi été inscrit sous diverses formes (projets semi-autonomes, hackathons, sprint de quelques jours...) dans les programmes pédagogiques de plusieurs formations (développement, design, etc.), afin de faire émerger des concepts et des prototypes. Pour l'équipe MesInfos, cela a signifié une mobilisation continue.

L'effort a également porté sur le développement de services par les partenaires détenteurs de données, chaque partenaire restituant des données étant encouragé à développer un niveau de visualisation "de base" de ses données, afin que les testeurs ne soient pas uniquement confrontés à des données "brutes", souvent peu compréhensibles. Des ateliers rassemblant les partenaires ont également été organisés afin de faire émerger des croisements de données fertiles.

Au-delà de ces chantiers animés sur la durée, **tout le monde peut participer au pilote en développant des services**. Pour soutenir cet effort, un "Kit développeur" a été constitué, rassemblant la documentation des données ainsi que des données de synthèse (données types) et un kit de développement à proprement parler.

En 2017, des challenges devraient être organisés afin de faire émerger d'autres prototypes.

Points de vigilance

Le chantier "réutilisation" n'est pas simple et il est complexe de motiver des développeurs, pour plusieurs raisons : le marché du Self Data est encore émergent, sans modèle économique stabilisé. On demande par ailleurs aux développeurs de développer pour une plateforme (Cozy) des applications propres à cette plateforme, avec un nombre d'utilisateurs réduits... la voie du cloud personnel et du Self Data est exigeante !

Une organisation qui restitue des données dans le cadre du pilote peut-elle également développer des services dans le pilote ?

Rien n'interdit à un détenteur de données d'être par ailleurs réutilisateur dans le cadre du pilote, bien au contraire. Il est important, à minima, que les acteurs qui restituent les données proposent une couche servicielle simple, permettant aux individus de les visualiser. Mais il est aussi possible de développer des services qui croiseront plusieurs données, sous le contrôle de l'individu.

Qui est responsable du traitement lorsque des données hébergées par un cloud personnel sont utilisées par un service tiers ? Quel consentement doit obtenir l'application ?

De manière assez classique, les fournisseurs des applications tierces sont responsables des traitements effectués sur les données utilisées par ces applications.

Lorsque l'individu installe une application, il donne son consentement pour que l'application accède à certaines données, pour certains usages (selon un principe de finalité). Le fournisseur de l'application est encouragé à faire un effort quant à la lisibilité des conditions générales d'utilisation et du recueil du consentement.

Suis-je propriétaire de mon service ? Puis-je monétiser mon service ? Mon code est-il ouvert / visible par tous ?

Oui, vous êtes propriétaire du service, vous pouvez le monétiser à votre guise, et vous pouvez protéger votre code comme il vous convient. Cependant, il faudra parfois imaginer des solutions techniques compatibles avec la plateforme, notamment ouvrir une partie de votre code.

Puis-je avoir accès aux données personnelles des utilisateurs de Cozy (sous-entendu qui utilisent mon service ?)

A priori non, car ce n'est pas nécessaire. Vous n'êtes pas obligés d'y avoir accès (vous aurez accès seulement à des données type/fictives vous permettant de développer votre service), pour offrir un service riche en fonctionnalités ! Cependant, les dispositions techniques ne suffisent pas seule à bloquer les intentions malveillantes. C'est pourquoi dans cette phase de démarrage, chaque application proposée aux testeurs est auditée avant d'être mise à disposition des testeurs.

VI. Recrutement du panel

Comment s'y est-on pris ?

Le recrutement des testeurs du pilote est assuré en deux phases par la Fing et la MAIF :

- un premier temps avec des beta-testeurs recrutés en grande partie parmi les collaborateurs du partenaire pilote du projet (Maif), et également grâce à un recrutement plus ouvert (permettant à de premiers curieux motivés de s'inscrire comme testeurs). À ce stade, l'ambition était d'avoir des beta-testeurs motivés, prêts à tester la plateforme et les services et à faire des retours ; nous avons décidé de ne pas communiquer sur ce recrutement. En tout, 300 personnes ont été recrutées.
- un second temps, plus large, incluant environ 1000 à 2500 testeurs de plus, également recrutés parmi les sociétaires Maif, prévu pour courant 2017.

Peu de critères ont été définis pour le recrutement de ce "panel", car nous ne recherchons pas la représentativité. Les seuls critères du recrutement sont la motivation et le fait d'être client d'au moins une (voire davantage) organisations partenaires restituant de la donnée, afin que les cas d'usages permis soient plus riches et divers (croisant davantage de données).

Ce recrutement en deux temps présente plusieurs avantages pour le pilote : la possibilité d'identifier et de corriger des bugs autour de la transmission des données, de la plateforme ou des services tiers ; la possibilité d'éprouver les outils de la recherche (questionnaires, etc.) ; cela laisse aussi du temps pour que des services tiers soient développés.

Enfin, il est à noter que les personnes recrutées ont le choix de participer ou non au projet de recherche, qui se penche sur l'appropriation et les usages des données, des services et du cloud personnel par les testeurs.

Une animation des testeurs, basée sur des sujets de discussions proposés sur un forum dédié ainsi que des newsletters, est mise en place. D'autres modalités présentiels (focus groups ou ateliers) pourront également être mises en place.

Early adopters ou recrutés aux usages numériques divers ?

Le recrutement est très ouvert, et ne s'adresse pas qu'aux personnes ayant des usages numériques très avancés. Néanmoins, nous avons tout de même choisi de privilégier des profils utilisant le numérique, puisque le recrutement lui-même s'opère en ligne.

Des déclarations de traitement spécifiques à faire ?

Le recrutement a nécessité de faire plusieurs déclarations de traitement : une déclaration simplifiée pour la constitution de la liste des participants et son traitement, et une déclaration pour la dimension spécifique recherche, permettant d'entrer en contact avec les testeurs par d'autres canaux (entretiens, traces d'usage, etc.)

Les organisations qui restituent les données sont-elles en relation avec leurs clients testeurs ?

A part dans le cas de la MAIF, les organisations détentrices de données ne recrutent pas leurs clients. Elles peuvent néanmoins entrer en relation avec leurs clients si elles souhaitent tester des idées ou obtenir des retours sur des services, notamment par le biais du forum mis en place pour le pilote.

VII. Observation et recherche

Comment s'y est-on pris ?

Le pilote MesInfos est le terrain d'un projet de recherche, visant essentiellement à comprendre les usages que les individus peuvent faire de leurs données, de leur cloud personnel et de services tiers.

L'équipe de recherche est composée de 5 chercheurs reconnus en sciences sociales et économiques, coordonnés par une personne dédiée aux sujets de l'animation des testeurs et de la recherche, au sein de l'équipe MesInfos.

La recherche s'appuie sur plusieurs méthodes :

- Des questionnaires pour connaître l'avis et les impressions concernant le Cozy (trois questionnaires en totalité, les réponses seront récoltées de manière anonymisée (1) ou à minima pseudonymisée (2), aucun nom/prénom ne sera demandé).
- Un suivi de l'activité des testeurs sur le Cozy Cloud (les chercheurs s'intéresseront uniquement à la fréquence d'utilisation du Cozy Cloud, ils n'auront accès à aucune donnée personnelle).
- Entretiens individuels/entretiens de groupe (les chercheurs inviteront les testeurs à participer à des entretiens afin d'approfondir leur réflexion sur la valeur d'usage du Cozy Cloud. La participation est volontaire et non obligatoire).

1) Effacement des données permettant l'identification d'une personne.

2) Remplacement des noms par des pseudonymes ne permettant pas d'identification de la personne.

Plusieurs phases d'observation et d'analyse sont prévues dans le pilote : une première phase, après quelques semaines d'utilisation du cloud personnel par les testeurs ; la deuxième et la troisième quelques mois plus tard.

Le recueil des informations par le biais des moyens décrits plus haut se fera de manière anonymisée ou à minima pseudonymisée. Les informations ne seront accessibles qu'aux chercheurs du pilote MesInfos qui ont signé la charte de confidentialité et de bonnes pratiques. De plus, il est important de noter que les informations ne servent qu'à des fins de recherche et aucunement à faire des offres commerciales, que les résultats des analyses et les livrables recherche pourront être partagés ; enfin, le pilote MesInfos dans son intégralité se déroule bien sûr sous la supervision de la CNIL.

Si la recherche a amorcé ses travaux en 2016 et au premier semestre 2017 (conception des outils, protocole de recherche, etc.), c'est à partir du second semestre 2017, une fois que les testeurs auront pris leur Cozy et leurs données en main que les chercheurs débiteront réellement.

Ce document est une première version publiée à mi-chemin, essentiellement rédigée au premier trimestre 2017. Elle sera remaniée et enrichie au fil du pilote, afin de documenter ses principaux enseignements.

Retrouvez les travaux du projet MesInfos sur le site : www.mesinfos.fing.org