

La portabilité des données en pratique

Saisir l'opportunité du nouveau droit à la portabilité des données personnelles.

Ce cahier porte sur le droit à la portabilité, instauré par le Règlement européen sur la protection des données (RGPD), qui entre en application le 25 mai 2018. Il défriche le sujet des données concernées par ce droit, et celui de sa mise en oeuvre au sein des organisations.

Ce cahier est le résultat d'un travail commun réalisé dans le cadre de deux projets aux expertises complémentaires de la Fing : celles de "Parlez-vous Data ?" en matière de diffusion d'une culture Data et celles de "MesInfos" en matière de Self Data.

Le groupe a été animé par Guillaume Jacquart (MesInfos), avec l'étroite collaboration d'Armelle Gilliard et Charles Népote (Parlez-vous Data ?). Le groupe était composé de Hugues Roumezin, François Herlent pour la Maif, Jean-Marc Raibaud pour Orange, Régis Chatellier pour la CNIL. Ont également participé à ce groupe de travail : Érik Arnaud pour la Maif, Emmanuel Malkor pour Carrefour, Pauline Kullmann pour GRDF, Amine Essafoui pour GS1. Le cahier a été rédigé sous la direction de Guillaume Jacquart.

Retrouver ce document en ligne :

<http://infolabs.io/dp17>



“Parlez-vous Data ?” - Infolab **“Le potentiel des données enfin partagé”**

Depuis 2013, la Fing explore des dispositifs d'acculturation aux données, pour mettre les acteurs en capacité d'exploiter les données dans leurs activités et dans leurs projets. En 2017, le programme opérationnel “Parlez-vous Data ?” explore le déploiement concret de la culture de la donnée dans les organisations, alternant travaux collectifs, partages de pratiques et déploiement opérationnel in situ.

<http://fing.org/?Infolab>



MesInfos

“Si j'ai une donnée sur vous, vous l'avez aussi. Et vous en faites ... ce qui a du sens pour vous !”

Depuis 2012, la Fing explore au sein du projet MesInfos la question du Self Data : “la production, l'exploitation et le partage de données personnelles par les individus, sous leur contrôle et à leurs propres fins”. Les lignes directrices du G29 sur le droit à la portabilité citent le projet MesInfos comme l'un des démonstrateurs pionniers du potentiel d'innovation que cette avancée législative représente.

<http://mesinfos.fing.org/>

INTRODUCTION

Le règlement européen sur la protection des données (RGPD) décrit dans son [article 20](#) un nouveau droit, qui “offre aux individus un chemin facile pour récupérer et réutiliser leurs données personnelles par eux-mêmes” : le droit à la portabilité. Les données doivent être mises à disposition facilement, “dans un format structuré, d'usage commun et lisible par des machines”. L'utilisateur peut vouloir les télécharger chez lui, ou les mettre directement à disposition d'un tiers.

À la différence du droit d'accès qui permet aux individus de vérifier ce que les organisations savent sur eux, et ainsi apporter un peu de transparence et de confiance, le nouveau droit à la portabilité tend à rééquilibrer la balance, entre responsable de traitement et personne concernée, sur la valeur qui peut être tirée des données. Les organisations ont développé des outils et collecté de plus en plus de données ces dernières années, pour en tirer une valeur importante. Le droit à la portabilité invite à ce que l'individu puisse lui aussi réellement accéder à ces données et en tirer des usages.

L'esprit est donc de proposer aux individus la même qualité de services sur les données dont les organisations disposent en interne. Ni plus - il ne s'agit pas de collecter de nouvelles données à l'attention des personnes concernées, ni moins - ces données doivent être utilisables par les individus, autant qu'elles sont utilisables dans les organisations. Récupérables sans démarches complexes, régulièrement, rapidement, à un coût proportionné (si cela s'avère nécessaire) et utilisables ...

Pour s'orienter dans cette nouvelle démarche, le groupe de travail s'est attelé à deux sujets : les données concernées par le droit, et la manière d'organiser un projet de mise en oeuvre de la portabilité.

Quelles données sont concernées par la portabilité ?

Contrairement au droit d'accès, le droit à la portabilité ne concerne qu'un **sous-ensemble des données personnelles** détenues par le responsable de traitement. Pour faciliter la délimitation du périmètre des données concernées par la portabilité, la CNIL a proposé un outil simple sous la forme de 5 questions.

Nous vous proposons de reprendre ces 5 questions en illustrant des cas limites et cas particuliers. Commençons par la règle.

Pour une donnée, si le responsable de traitement répond :

- "oui" aux 5 questions, alors la donnée est portable ;
- "non" à au moins une question, la donnée n'est pas portable ;
- "oui mais" à certaines questions, alors la donnée est sans doute portable, mais demande un peu de travail. Plusieurs cas particuliers sont précisés dans la suite du document.

1 - Les données personnelles concernent-elles la personne ?

☒ oui ☐ non

- Oui mais : si l'utilisateur du service n'est pas le contractant (ex : téléphone mobile, boîte email, ...)
- Oui mais : le service est collectif, il peut-être utilisé par plusieurs personnes (simultanément ou non).
- Oui mais : la personne concernée est difficile à discerner automatiquement.

2 - Le traitement est-il effectué par des moyens automatisés ?

☒ oui ☐ non

- Oui : les traitements papiers ne sont pas concernés par la portabilité.
- Non : même les données archivées (difficiles d'accès). Elles peuvent seulement justifier un délai de restitution plus long (mais d'autre part, est-ce légitime de conserver ces données aussi longtemps ?).

3 - La base légale de la collecte est-elle le consentement ou le contrat ?

☒ oui ☐ non

- Consentement = Oui
- Contrat = Oui
- Obligation légale = Non
- Sauvegarde des intérêts vitaux = Non
- Mission d'intérêt public = Non
- Intérêt légitime = Non

4 - Les données sont-elles fournies par la personne concernée ?

☒ oui ☐ non

- Oui : Toutes les données explicitement transmises par l'individu.
- Oui : Également, toutes les données résultant d'actions directes de la part de l'individu (transactions, traces d'usage, ...)
- Non : Cela exclut (du périmètre obligatoire) les données issues du profilage par exemple.
- Mais OUI ! il faut simplement s'intéresser aux données réutilisables plutôt qu'aux seules données portables.

5 - Les données portées ne nuiront pas aux droits et libertés d'autrui ?

☒ oui ☐ non

- Oui mais : la donnée respecte-t-elle la réglementation vis-à-vis des salariés ?
- Oui mais : la donnée concerne plusieurs personnes ?

Oui mais : les cas difficiles

1 - Les données personnelles concernent-elles la personne ?

L'utilisateur du service n'est pas le contractant, le service est collectif, le service peut-être utilisé par plusieurs personnes, la donnée concerne plusieurs personnes ...

Quelques exemples pour mettre en valeur chaque cas :

L'utilisateur du service n'est pas le contractant : l'historique d'appel pour un téléphone portable dont le contrat est au nom des parents de la personne concernée.

Le service peut-être utilisé par plusieurs personnes : la carte de fidélité au supermarché.

Le problème dans tous ces cas, est que la personne concernée n'est pas clairement identifiée au sein du système d'information de l'organisation. Il s'agit tout de même de données personnelles : la personne concernée peut-être ré-identifiée grâce à ces données ; mais qui ne sont pas attribuables automatiquement dans l'état actuel du service.

La portabilité amène de nouvelles contraintes, qui parfois remettent en question des choix sur les modèles de contrats, les modèles de données et même les processus dans le monde réel des services. Ces situations, qui n'étaient pas traitées rigoureusement jusqu'alors mais qui fonctionnaient, deviendront non-conformes avec le RGPD, et pénaliseront également la qualité et la réutilisabilité des données. Elles nécessiteront la mise en oeuvre de projets importants pour parvenir à une situation pleinement satisfaisante.

5 - Les données portées ne nuiront pas aux droits et libertés d'autrui ?

La portabilité est un droit, mais il n'en est pas pour autant prioritaire sur d'autres droits, comme le respect de la vie privée par exemple. Certaines données, certains cas de figure forceront à des compromis entre l'application du droit à la portabilité et le respect d'autres droits. Néanmoins, ce n'est pas une démarche nouvelle. La question se pose déjà sur le droit d'accès. Mais aussi, dans une situation proche de la portabilité en termes de volume : les restitutions "papier" d'informations personnelles (comme les relevés détaillés d'appel téléphonique) font déjà l'objet de nombreuses vérifications et annotations avant d'être transmises au client. Cependant, la quantité et la diversité des données laissent présager beaucoup de travail pour identifier, anticiper et trouver des parades pour ces cas particuliers.

Oui mais ne signifie pas NON

Quelles que soient les difficultés techniques ou organisationnelles que pose la mise en conformité, ainsi que son apparente nouveauté, il s'agit d'un droit des utilisateurs. La solution ne peut donc être "je ne sais pas comment rendre ma donnée portable, alors je continue comme avant".

Lorsqu'une organisation demande, via un contrat ou un consentement, à effectuer un traitement sur les données personnelles d'un individu, elle en prend également les responsabilités, notam-

ment celle de se conformer au droit à la portabilité. Si elle n'en est pas capable, rien ne l'oblige à traiter ces données. Il lui suffit donc de purger ces données-là pour être en conformité.

Portables et réutilisables

Pourquoi un nouveau droit ?

RGPD - Article 20 - Droit à la portabilité des données : *“Les personnes concernées ont le droit de recevoir [...] dans un format structuré, couramment utilisé et lisible par machine, et ont le droit de transmettre ces données à un autre responsable de traitement ...”*

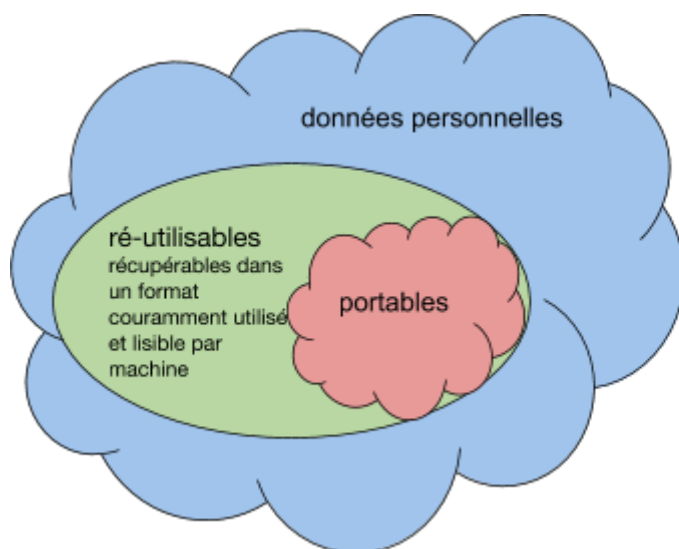
Recommandations du G29¹: *“Ce nouveau droit a pour objectif de responsabiliser les personnes concernées et de leur permettre de contrôler da-*

vantage les données à caractère personnel les concernant. [...] Il facilitera le passage d'un prestataire de services à un autre et encouragera dès lors la mise au point de nouveaux services dans le contexte de la stratégie pour un marché unique numérique..”

La portabilité est un nouveau droit pour les individus. Il contribue à rééquilibrer les forces entre organisation et individus. Il crée également, comme le souligne le G29, une opportunité de proposer de **nouveaux services**, ou de se **différencier par rapport à la concurrence** pour les entreprises.

Le but poursuivi ici est donc de suivre l'impulsion donnée par le droit à la portabilité, en décrivant comment les données que l'organisation utilise tous les jours peuvent être partagées avec les individus qu'elles concernent, afin qu'elles puissent leur simplifier la vie à eux aussi ! On cherchera donc **comment on peut rendre les données réutilisables**, et non seulement à se prémunir de difficultés judiciaires.

ILLUSTRATION



1. Lignes directrices relatives au droit à la portabilité des données. G29

On peut appliquer cet objectif de réutilisation d'une part au **périmètre** des données concernées : pour être en conformité, un responsable de traitement doit proposer un ensemble de données réutilisable, qui doit inclure les données personnelles dans ce périmètre, mais rien ne l'empêche, de proposer un périmètre plus large ! Et d'autre part, à leur format, leur documentation ... Sont-elles réellement intelligibles, **réutilisables** par un autre responsable de traitement ?

Réutilisabilité pour les personnes concernées : pour quoi faire ?

“Notre première question a d'abord été la suivante : quelle va être l'utilisation possible de ces données par le sociétaire ?”. C'était prioritaire à celle de l'obligation légale” (MAIF)

Les trois possibilités de réutilisation suivantes invitent à regarder les données de son organisation sous un angle nouveau, dans des usages différents, et actionnées par les personnes concernées. C'est un exercice qui doit permettre de transformer le “Ma” de “Ma donnée” en tant qu'organisation, outils de travail quotidien ; en “Ma donnée” en tant que personne concernée (individu, client, citoyen ...), qui pourrait me rendre de multiples services !

Même si le but ici n'est pas de fixer à l'avance les nouveaux usages qui seront faits des données, car la plupart d'entre eux seront découverts et connus seulement après leur mise à disposition, cet exercice permet néanmoins de dresser des priorités, et aussi de se projeter sur les méta-données, les informations complémentaires et les outils qui devront accompagner les données pour que leur réutilisation soit possible.

Ma donnée correspond à l'un des grands cas d'usage de la relation client.

L'historique de la relation, la traçabilité des opérations, ... : ces données sont utiles à l'organisation, lui permettent d'être efficace dans ses échanges avec ses usagers. De la même manière, ces données pourraient être utilisées par les usagers

pour leur faciliter le suivi de leurs dossiers en cours avec des fournisseurs (pièces déjà communiquées, numéros de dossiers, identifications des interlocuteurs précédents, ...).

Ma donnée est utilisée dans d'autres usages avérés.

Exemples :

- vérifier ses comptes, faire son budget à partir de l'historique de ses opérations bancaires,
- les données de consommation permettent au client de faire une analyse d'impact de ses travaux, savoir si un nouveau venu dans le foyer fait une différence, si les gestes “développement durable” permettent de faire des économies...
- ...

Ma donnée ouvre à de nouveaux usages, palpables, pour l'individu.

- Mes données de géolocalisation me permettent de reconstituer mon agenda, mon histoire, éventuellement mes goûts, etc.
- La durée passée au téléphone me permet de savoir rétroactivement quand et pour quelle durée s'est effectué un appel.
- ...

> Aller plus loin : [*“Mais que peut-on faire à partir de ses propres données ?” - Schéma des 7 domaines d'usage. \(MesInfos\)*](#)

Réutilisabilité, est-ce que ma donnée a un sens ?

Ma donnée personnelle a-t-elle un sens à être partagée au client ? À tous les clients ?

- Des données trop anciennes ? (il faut regarder de plus près les limites de conservation de ces données, à ne pas confondre avec les données qui doivent être historisées mais non exploitées).
- Des données dont la qualité n'est pas suffisante...

Deux stratégies pour ces cas :

1. On peut raisonner sous la forme : “je ne suis pas sûr que ça intéresse mon client MAIS ça ne me coûte presque rien donc dans le doute, je les lui partage”.
2. Ou bien, “ces données amènent plus de problèmes que de solutions, et elles seront ajoutées à la liste des données purgées pour la mise en conformité avec le RGDP.”

Réutilisabilité : les données sont-elles intelligibles ?

Ma donnée personnelle est-elle immédiatement lisible (intelligible) pour l'individu ?

Ce n'est pas forcément l'objectif recherché dans le contexte de la portabilité. En effet, si avec le droit d'accès, c'est bien la lisibilité des informations dont dispose le responsable de traitement qui est recherché ; avec la portabilité, l'objectif est plutôt de donner de nouvelles capacités, de redonner du pouvoir aux individus, au moyen des données qui les concernent, sans doute via des outils dédiés. La valeur des données sera donc présente dans leur capacité à lever le blocus sur des usages pour les individus, et non dans la simple relecture de ces informations. Ainsi, plutôt que de chercher la lisibilité, on cherchera à rendre ses données interprétables en vue d'être réutilisées, notamment via des tiers.

Ma donnée personnelle est-elle interprétable, en dehors de l'organisation ?

- La donnée personnelle est une référence dans une autre base de données (non personnelles), interne de l'entreprise. Exemple : ma playlist Spotify : c'est une liste des références des morceaux de musique, dans le système d'information de Spotify.
- La donnée personnelle fait référence au nom d'un produit, d'une formule commerciale, ... Exemple : contrat : *Raqvam*, ou encore, *options : Base*.

Dans ces deux cas, la donnée n'est pas exploi-

table sans des éléments de contexte, qui ne sont souvent pas publics.

Pour le premier cas des références techniques, il faudrait, soit ouvrir (Open Data) la table de référence, soit enrichir les données avec des données plus facilement exploitables (pour l'exemple de la playlist Spotify : nom de la piste, artiste, date, ...), ou encore, proposer un identifiant dans un référentiel public, s'il en existe pour le domaine (par exemple les codes barres dans [OpenFoodFacts](#)).

Le second cas est plus délicat car il conduit à des faux-sens : l'option “base” aura par exemple le même nom, mais pas les mêmes caractéristiques d'un concurrent à l'autre. C'est là où l'explicitation et la documentation des données prennent tout leur sens. Le risque étant de causer incompréhension, frustration et mécontentement chez les usagers lors de réutilisations futures de ces données.

Ma donnée personnelle, comment puis-je en faciliter la réutilisation ?

- Par exemple, si j'utilise le code INSEE de la commune pour définir un lieu, n'ai-je pas intérêt à fournir le nom de la commune en clair, afin de simplifier la réutilisation ?
- Si la donnée est conservée en interne sous forme de code (cf. paragraphe précédent), dois-je la remplacer par sa signification dans les données restituées ?
- Quel format, quel modèle pour ces données personnelles ?

Pour les deux premiers points, des données plus intelligibles au premier coup d'oeil seront en effet plus faciles à prendre en main et à réutiliser. Cependant, attention à ne pas appauvrir les données en leur donnant une forme plus intelligible. Par exemple avec les codes INSEE, 6 communes s'appellent Montréal en France, 26 communes Saint-Martin ... Attention également à ne pas alourdir les jeux de donnée en les noyant dans des explications.

Comme évoqué dans la première partie, les données, pour certaines dédiées à des usages très précis au sein de l'organisation, peuvent être utilisées pour des usages très divers. Pour

le format et le modèle choisi pour les restituer, deux cas de figure : **soit** il existe un (des) **standard** largement utilisé pour ce type de données (et pour se faire une première idée, on pourra s'inspirer par exemple des choix de Google dans son service [takeout](#)). **Soit** il n'en existe pas (encore). Le plus simple est alors sans doute de rester **au plus proche du modèle utilisé en interne**. Quitte à apporter des modifications profondes dans quelques années, quand des usages prépondérants sur la réutilisation de ces données se seront distingués ; ou bien quand un standard aura émergé. Et si la préexistence de standard n'est pas absolument nécessaire, c'est un sujet à initier ou à rejoindre rapidement. D'une part parce que les standards vont aider à débloquer la valeur des données et d'autre part pour s'assurer de ne pas subir le standard imposé par des concurrents ou de nouveaux entrants.

Comment organiser un projet de portabilité au sein d'une organisation ?

Un projet transverse

La mise en oeuvre de la portabilité demande la collaboration de plusieurs services de l'organisation : juridique (dont le DPO), Systèmes d'information, gouvernance de la donnée, et "top management", vu les risques encourus en termes d'image et désormais de sanctions financières. Pour faire de la portabilité une opportunité, d'autres services doivent être impliqués : les différents métiers qui produisent et utilisent ces données au quotidien. Ce sont, dans de nombreux cas les seuls à avoir une connaissance précise et complète des données, à même de maintenir une documentation et de répondre à

des questions sur les données. Également les services de communication, tant en interne, pour faire connaître le projet et la démarche et permettre au projet d'avancer ; qu'en externe afin de valoriser cette démarche et ainsi recueillir le fruit de ce travail sous forme d'utilisateurs satisfaits, fidélisés.

Un projet tourné vers l'extérieur

La portabilité exige de faire sortir de l'organisation - plus précisément de ses systèmes d'information - des données qui y étaient jusqu'à présent soigneusement verrouillées. Cela signifie d'une part, se mettre dans la position de **se rendre intelligible** à l'extérieur de l'entreprise, par exemple en traduisant les éléments de langage internes.

Mais surtout, cela implique de **se confronter** au plus tôt à ces interlocuteurs extérieurs : services tiers, réutilisateurs, utilisateurs. C'est en effet le moyen le plus rapide pour déverrouiller la valeur d'usage de ces données, mais aussi d'en découvrir leurs potentiels d'utilisages et favoriser le développement de nouveaux services qui pourraient naître, et promouvoir les autres données, personnelles ou non, publiques ou non, issues d'autres secteurs d'activité ou de concurrents, ... qui pourront inspirer l'innovation.

Un projet qu'il faut présenter et expliquer en interne

Comme présenté ci-dessus, ce projet implique la collaboration de multiples compétences au sein de l'entreprise. Mais surtout, c'est une démarche qui apparaît en **opposition** avec les principes de bases appliqués depuis de nombreuses années au sein des entreprises. *"Faire sortir" des données personnelles, si "précieuses", "complexes" et "sensibles", pour les mettre dans les mains des individus qu'elles concernent : qui ne savent pas qu'elles existent, ne veulent pas le savoir et de toute façon ne pourront rien en faire à part se les faire voler ou les donner bêatement au concurrent direct : c'est tout simplement une hérésie !"*

Au-delà de cette première réaction, cette démarche amène des questionnements légitimes à chaque collaborateur impliqué dans le projet. Les risques de la mise en oeuvre de la portabilité sont bien discernables : ce sont ceux qui sont combattus au quotidien - risques de sécurité,... - par les organisations, avec une complexité accrue. Et comme la démarche s'écarte des habitudes, c'est une prise de risque qui semble individuelle.

Au contraire, les opportunités qui s'ouvrent auront des effets sur un temps plus long et avec des retombées beaucoup plus diffuses pour l'organisation (et aussi chaque individu, l'écosystème d'innovation, ...). En tout cas pas individuelles. C'est bien un choix, un virage et un investissement **de long terme** de l'organisation qui doit être **expliqué comme tel**. L'aspect vertueux et sociétal du projet n'est pas à négliger, chaque collaborateur étant aussi usager !

Ce sont ces facteurs qui ont forgé la proposition de schématisation de mise en oeuvre du droit à la portabilité ci-dessous. Cette proposition est issue d'un travail collectif réunissant des participants de grandes organisations, et se base sur un véritable travail de terrain effectué au cours du projet MesInfos, dans lequel des grandes organisations font l'effort de restituer à des testeurs les données qu'elles ont sur eux (première expérimentation en 2013 et pilote MesInfos à partir de 2016).



Par les participants du groupe de travail
portabilité, de *Parlez-vous data ?* et
Meslins de la Fing, réunissant la *Maif*,
Orange, la *CNIL*, *Carrefour*, *CRDF*, sous la
direction de *Guillaume Jacquart*, *Charles*
Nepote et *Armelle Gillard*.

Élargissement à des acteurs
extérieurs à l'organisation



Durée

3 Mois

2 Semaines

Acteurs

Communication	Services tiers
Usagers	Top Managers
Métiers	Gouv. données
DSI	Juriste DPO

Licence



<http://info.labs.io/schema-portabilite>

LA PORTABILITÉ EN PRATIQUE

Élargissement à des acteurs extérieurs
à l'organisation



Mise 2
en oeuvre

1 Préparation

Accompagnement
Apprentissage 3

PROJET

PROJET

PROJET

PROJET

PROJET

Présenter la
portabilité

Proposer
une liste de
données
éligibles

Analyse de
la liste >
Validation

Choix des
dispositifs de
transmission

Transmission

Documentation

Acculturation
des individus

Animation et
participation
à la
ré-utilisation

Contribuer à
l'écosystème
et en tirer
partie



PRÉSENTATION
EN INTERNE
DE L'OPPORTUNITÉ

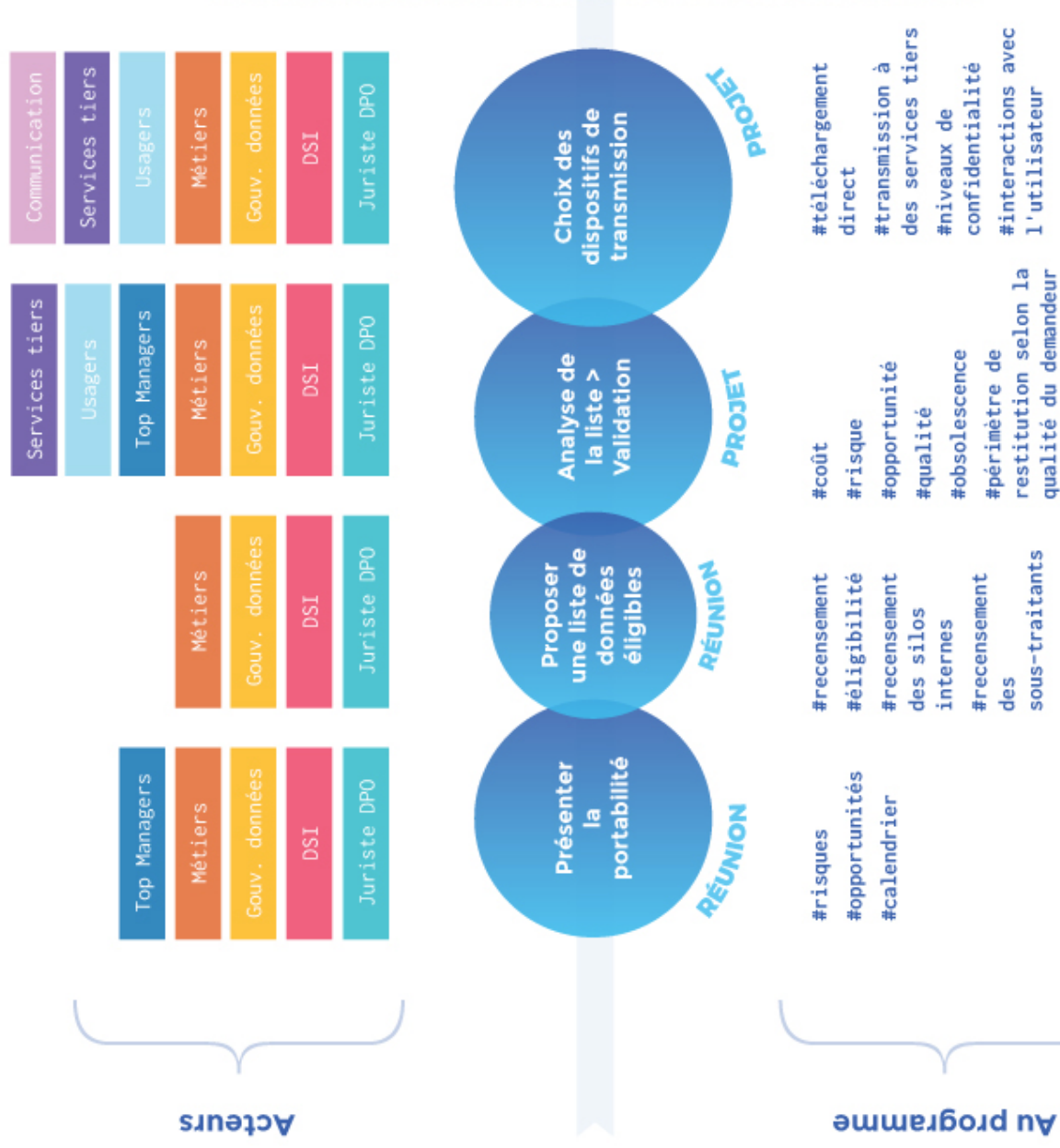
Opportunité pour chacun - risques - calendrier

Élargissement à des acteurs extérieurs à l'organisation



1. Préparation

En détail



PRÉSENTATION EN INTERNE DE L'OPPORTUNITÉ

Opportunité pour chacun - risques - calendrier

Étapes

- 1 Préparation
- 2 Mise en oeuvre
- 3 Accompagnement Apprentissage

Durée

2 Semaines

+

3 Mois

-

Élargissement à des acteurs extérieurs à l'organisation

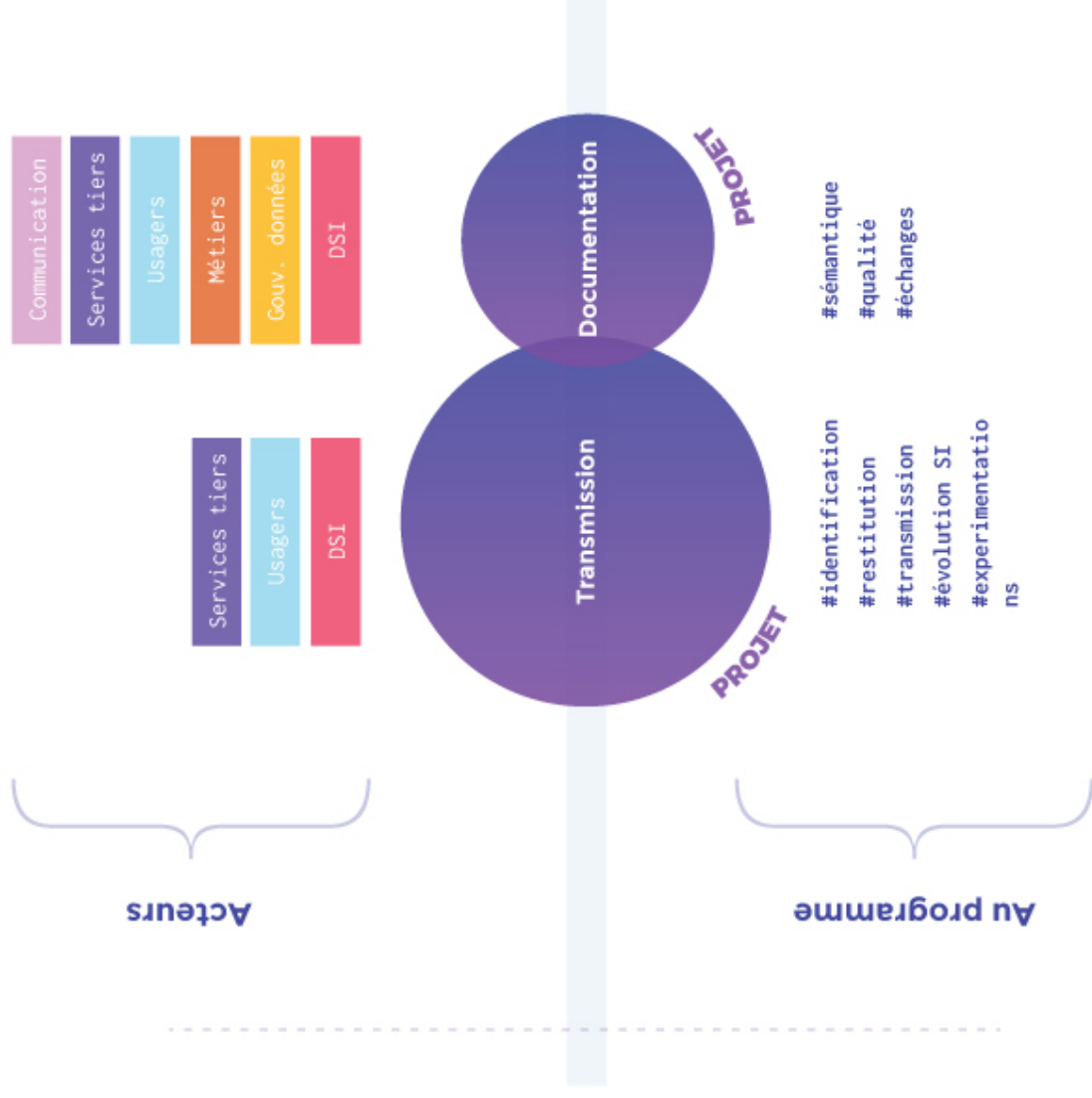


Licence



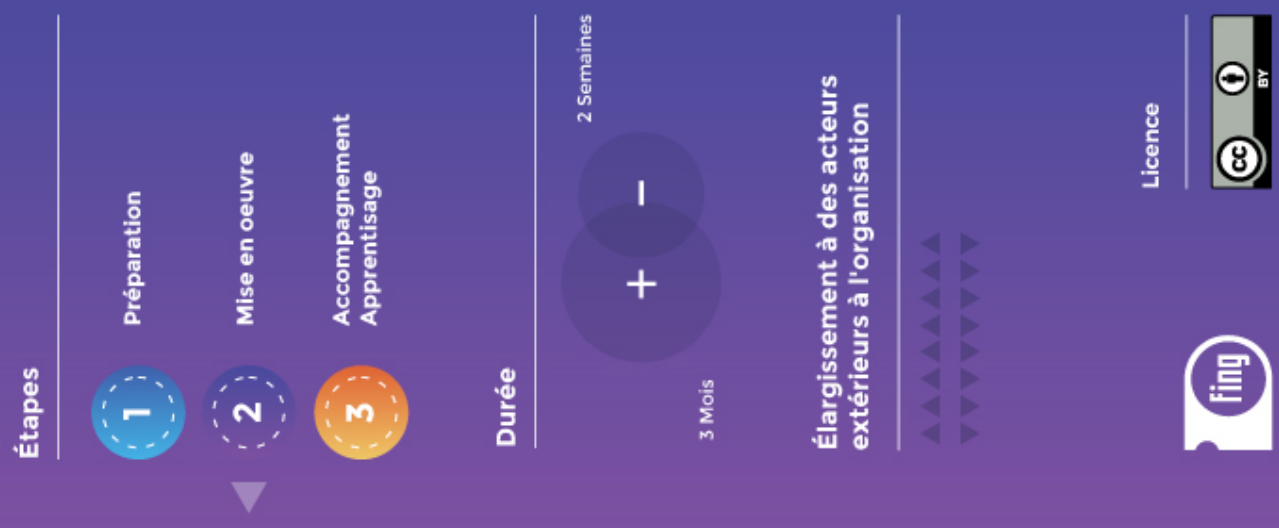
2. Mise en oeuvre

En détail



PRÉSENTATION EN INTERNE DE L'OPPORTUNITÉ

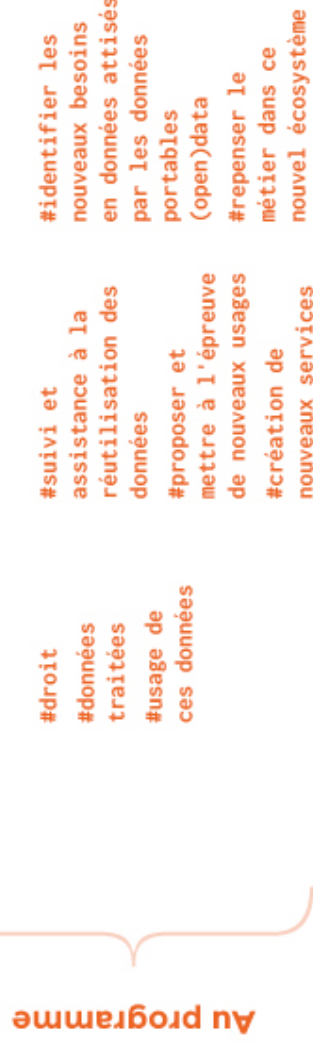
Opportunité pour chacun - risques - calendrier



3. Accompagnement Apprentissage

En détail

Élargissement à des acteurs extérieurs à l'organisation



Élargissement à des acteurs extérieurs à l'organisation

Étapes

1 Préparation

2 Mise en oeuvre

3 Accompagnement Apprentissage

Durée

2 Semaines

+

3 Mois

Élargissement à des acteurs extérieurs à l'organisation



Licence



CAHIER #4

La portabilité des données en pratique

<http://infolabs.io/dp17>



CC License Creative Commons BY 4.0

Ce cahier est une production de la "Campagne Infolab - Parlez-vous Data ?" et du "Projet MesInfos" de la Fing



avec le soutien des partenaires MesInfos et Parlez-vous Data ?



La Fing a le soutien de ses grands partenaires

