

LA MESURE DE L'IMPACT DES DONNÉES OUVERTES PAR L'USAGE

Sommaire

Introduction	4
Partie I : Impact, usages et réutilisations : enjeux de définition	5
I. Le défi de l'évaluation des impacts de l'open data	6
II. Comment définir l'impact de l'open data ?	7
L'OURdata Index de l'OCDE	7
L'Open Data Maturity Index de la Commission européenne	8
L'Open Data Barometer	8
L'Association pour la Fondation d'un Internet Nouvelle Génération (FING)	9
Les niveaux d'impact	10
Le rapport "Pour une politique de la donnée, des algorithmes et des codes sources"	11
III. Comment définir une réutilisation de données ?	11
Pourquoi s'intéresser aux réutilisations ?	12
Partie II : Comprendre les usages et les mécanismes de leur référencement	13
I. Analyse du catalogue des réutilisations sur data.gouv.fr	14
Analyse quantitative	14
Analyse qualitative	19
II. Identification de communautés de réutilisateurs	19
Etablir une typologie de communautés	19
Mécanismes de référencement des réutilisations	21
Partie III : Suivi et connaissance des usages par les producteurs de données	22

I. Méthodologie	23
II. Etat des lieux	24
III. Enjeux de la connaissance des usages	25
IV. Leviers méthodologiques et techniques	26
V. Besoins	27
VI. Perspectives et opportunités	27
Panorama des usages	28
I. Usages sectoriels	29
Usages par l'administration : l'exemple des startups d'État	29
Usages dans le secteur des mobilités et des transports	30
Usages dans l'enseignement supérieur	32
II. Usages du service public de la donnée	33
Définition	33
Etat de lieux des usages sur data.gouv.fr	33
Cas d'usage de réutilisation de données ouvertes	38
Conclusion	48
Accompagner les producteurs et réutilisateurs de données	48
Améliorer l'expérience des réutilisations de données sur data.gouv.fr	50
Bibliographie sélective	52
Études et rapports	52
Articles scientifiques	54
Ouvrages fondamentaux	56
Autres articles et ressources	56
Annexes	57
Analyse des réutilisations sur data.gouv.fr	
du Service Public de la Donnée	58
Benchmark de documentation de cas d'usage	60

INTRODUCTION

"Pour les citoyens, un enjeu de transparence de la vie publique et de simplification des démarches [...], pour les entreprises, susciter l'innovation et moderniser l'économie [...]; pour les acteurs publics, éclairer la décision et améliorer la qualité des services publics [...]; pour la recherche, fertiliser l'activité et favoriser les collaborations" : les impacts de l'ouverture des données publiques ont été largement documentés et justifient la mise en œuvre d'une politique ambitieuse en la matière.

Pourtant, plus de dix ans après les premiers mouvements d'ouverture, les mesures d'impact de l'open data sont ponctuelles. Ainsi, le rapport *"Pour une politique publique des données, des algorithmes et des codes sources"* recommande-t-il de *"conduire une évaluation de l'impact économique, social et scientifique de l'ouverture et du partage de données"*.

Pourquoi ces évaluations restent-elles encore ponctuelles, malgré un potentiel d'impact manifeste ? Comment mesurer l'impact de l'ouverture des données ? Quelles sont les réelles retombées de cette politique ?

Dans le cadre de nos différentes missions (conception et mise en œuvre de la politique publique de la donnée, accompagnement des producteurs de données et des réutilisateurs, etc.), nous avons entrepris de travailler sur ces questions et d'approfondir la mesure d'impact de l'open data. Le présent document de travail réunit l'ensemble des analyses et des enseignements collectés à date.

Une première partie s'intéressera à la définition des termes "impact" et "réutilisation" avec la revue de différents cadres méthodologiques (partie I). Il s'agira ensuite d'analyser les usages des données à travers le prisme des réutilisateurs de données (partie II) et des producteurs de données (partie III) avant de dresser un panorama de ces usages (partie IV). La conclusion présentera des recommandations et des perspectives, notamment pour la plateforme nationale de données publiques data.gouv.fr.

Le présent
document est
un document de
travail

PARTIE I : IMPACT, USAGES ET RÉUTILISATIONS : ENJEUX DE DÉFINITION

Les termes "impact" et "réutilisation" désignent des concepts à géométrie variable.

La première partie de cette étude sera consacrée à un travail de définition, à travers la revue de différents modèles, qui clarifiera le cadre de l'analyse qui suit.

I. LE DÉFI DE L'ÉVALUATION DES IMPACTS DE L'OPEN DATA

La question de la mesure des impacts de l'open data sur la société, sur l'élaboration et la mise en œuvre des politiques publiques ? et l'économie reste un sujet complexe auquel se confrontent toutes les administrations et organisations qui mènent une politique d'ouverture des données.

Plusieurs facteurs sous-tendent cette complexité : premièrement, un des principes fondateurs de l'open data est d'adopter un positionnement selon lequel il ne faut pas préjuger ni contrôler les usages futurs qui seront réalisés à partir des données mises à disposition.

Ensuite, la tâche d'évaluation exhaustive des impacts directs et indirects de l'ensemble des retombées (sociales, politiques, économiques, etc.) représenterait un effort conséquent et complexe voire équivalent à la publication des données elles-mêmes et demeure relativement peu mise en œuvre.

La complexité peut aussi provenir de la difficulté à définir ce qui entre dans le périmètre de la mesure d'impact. Les indicateurs les plus souvent utilisés pour mesurer la performance d'une politique d'ouverture des données ou d'un portail de données peuvent parfois être assimilés à des [indicateurs de vanité](#) en cela qu'ils ne permettent pas réellement de mesurer l'impact. Le nombre de visites sur un jeu de données, le nombre de téléchargements, etc. ne nous permettent pas de réellement savoir si l'ouverture a produit un impact concret même s'ils restent des indicateurs dignes d'intérêt. Il s'agit alors d'aller plus loin en mesurant plus systématiquement des indicateurs qualitatifs.

II. COMMENT DÉFINIR L'IMPACT DE L'OPEN DATA ?

Avec la montée en compétences des administrations sur le numérique, la mise en œuvre croissante du cadre juridique et la volonté des citoyens d'obtenir plus de transparence sur les politiques publiques, les données ouvertes sont devenues un sujet incontournable et ont un impact considérable sur la société. Aujourd'hui, chaque événement d'actualité majeure est accompagné d'une forte demande des citoyens pour la publication de données sur le sujet. Tous les secteurs sont concernés : économie, situation politique, environnement, santé, etc.

Il n'existe pas à proprement parler de définition réglementaire et/ou académique de l'impact des données ouvertes en France qui fasse consensus malgré les diverses entreprises en la matière.

Les impacts des données ouvertes sont multiples et peuvent être mesurés de différentes manières. La présente étude établit un rapide tour d'horizon des différents cadres méthodologiques.

L'OURdata Index de l'OCDE

L'OURdata Index est un indice qui évalue l'efficacité des politiques gouvernementales d'ouverture des données. Il s'agit de mesurer les efforts que font les gouvernements pour faciliter l'utilisation des données ouvertes. Il repose sur trois indicateurs :

- la disponibilité des données (notamment le nombre de jeux de données disponibles) ;
- l'accessibilité des données (notamment la qualité et la complétude des données) ;
- le soutien gouvernemental à la réutilisation des données (la promotion de l'open data, la mise en place de formations et la mesure de l'impact de l'ouverture).

Le troisième indicateur évalue dans quelle mesure les gouvernements jouent un rôle proactif dans la promotion de la réutilisation des données au sein et en dehors de l'administration. Il couvre principalement le principe 5 ("Amélioration de la gouvernance et de l'engagement des citoyens») et le principe 6 («Développement inclusif et innovation») de la Charte internationale des données ouvertes.

Il se décompose lui-même en trois sous-indicateurs : les initiatives et les partenariats de promotion des données, les programmes d'initiation aux données au sein du gouvernement et le suivi de l'impact.

Ce dernier examine l'évaluation de l'impact de l'ouverture des données par les gouvernements : recherches conduites ou financées, promotion d'initiatives de réutilisations des données ouvertes, etc.

L'Open Data Maturity Index de la Commission européenne

L'Open Data Maturity Index est un indice qui évalue les pays européens selon quatre critères :

- maturité des politiques publiques relatives à l'ouverture (*Policy*);
- qualité de la mesure de l'impact de l'open data (*Impact*);
- maturité du portail national d'open data (*Portal*);
- qualité des données ouvertes (*Quality*).

Sur la mesure de l'impact, la Commission européenne évalue notamment les mécanismes de suivi, ainsi que certains indicateurs comme le degré de réutilisation des données ouvertes. Il s'agit d'identifier

pour chaque pays des exemples de réutilisations dans quatre domaines : politique, sociale, environnementale et économique. Bien que la dimension «Impact» mette l'accent sur les cas de réutilisation des données ouvertes, la Commission européenne ne considère pas la disponibilité d'exemples de réutilisation comme une preuve directe de l'impact : l'impact de ces réutilisations doit également être mesuré.

L'Open Data Barometer

L'Open Data Barometer attribue un score à l'impact des données ouvertes de la trentaine de pays étudiés.

La mesure d'impact est réalisée selon six axes thématiques permettant une comparaison entre les pays :

- efficacité de l'action du gouvernement et des administrations
- transparence de l'action publique gouvernementale
- environnement / durabilité environnementale
- inclusion des groupes marginalisés dans la construction des politiques publiques et les processus décisionnels
- impact positif sur l'économie
- hauteur à laquelle les entrepreneurs s'emparent des données pour construire des entreprises / modèles économiques

L'Association pour la Fondation d'un Internet Nouvelle Génération (FING)

Dans son [rapport](#) "10 ans d'ouverture des données publiques", la FING (2000-2022) identifie cinq grandes familles d'impact de l'ouverture des données publiques, avec autant de questions ouvertes :

- les impacts économiques : l'open data a-t-il fait émerger un marché de la réutilisation des données ? Peut-on évaluer quantitativement ce marché ? Quels sont les acteurs économiques privés qui se sont emparés du sujet ?
- les impacts démocratiques : l'open data a-t-il renforcé la transparence de la vie publique ? Est-ce un outil de surveillance des acteurs de la vie publique (élus, agents des administrations, ...) et/ou un outil de capacitation des citoyens ?
- les impacts sur les organisations qui ont ouvert des données ou les utilisent, et leurs relations avec des tiers : l'open data a-t-il permis de nouer des relations avec des écosystèmes nouveaux ? En quoi les données ouvertes modifient-elles les relations existantes, par exemple avec les délégataires de service public ?
- Les impacts sur les métiers, les fonctions et les méthodes de travail : quels sont les métiers principalement impactés par l'ouverture des données ? Voit-on émerger de nouvelles fonctions, voire de nouveaux métiers ? L'open data a-t-il modifié les méthodes de travail au sein des organisations ?
- les impacts sur la diffusion d'une culture de la donnée au sein des organisations: l'open data comme brique première de la littératie des données ? La donnée ouverte a-t-elle ouvert la voie à des sujets comme la ville intelligente ?

Les niveaux d'impact

Etienne Pichot Damon a construit une [schématisation des indicateurs possibles](#) pour mesurer l'impact des données ouvertes à différents stades de leur utilisation.

Ces derniers se diversifient et changent en fonction du degré d'éloignement avec la donnée brute :

- Un premier niveau d'analyse se concentre sur la donnée brute elle-même : combien de visites et de téléchargement sur un jeu de données par exemple.
- Un second niveau d'analyse concerne les utilisations directes des données brutes : combien de réutilisations sont connues sur un jeu de données par exemple.
- Le troisième niveau d'analyse est celui des usages des réutilisations de données. Il s'agit par exemple du nombre d'utilisateurs d'une application utilisant des données ouvertes. Ce niveau devient plus difficile à documenter car nous n'avons pas accès aux audiences des différentes ressources produites à partir des données.
- Quant aux externalités (quatrième niveau d'analyse), qui désignent les impacts sociétaux des données ouvertes, elles ne peuvent être examinées qu'à de grandes échelles et sur des temps longs. A ce niveau, il est difficile d'attribuer à la seule donnée ouverte des modifications sociétales majeures comme des découvertes scientifiques ou la réorientation des politiques publiques. Il est également complexe de savoir si ces bouleversements sociétaux auraient eu lieu (et comment) en l'absence de données, la comparaison avec une situation "témoin" où les données ne seraient pas ouvertes étant impossible.

Le rapport "Pour une politique de la donnée, des algorithmes et des codes sources"

La mission Open Data conduite par le député Éric Bothorel définit quatre domaines d'impact de l'open data : économique comme moteur de l'innovation, démocratique pour améliorer le service public, politique pour restaurer la confiance des citoyens dans l'action publique et scientifique comme vecteur de connaissance.

[Le rapport sur la politique publique de la donnée, des algorithmes et des codes sources](#) issu de la mission recommande de mesurer davantage l'impact et constate une hausse de l'intérêt des producteurs de données concernant l'usage qui est fait de leurs données et la mesure de l'impact de l'ouverture.

III. COMMENT DÉFINIR UNE RÉUTILISATION DE DONNÉES ?

Il n'existe pas de définition stricte du concept de réutilisation de données publiques.

Une définition large veut qu'une réutilisation désigne l'utilisation des données publiques par des tiers à d'autres fins que celle de la mission de service public pour laquelle les données ont été produites ou reçues. Cette définition est celle retenue par la directive PSI (*"Les organismes du secteur public recueillent, produisent, reproduisent et diffusent des documents en vue d'accomplir leurs missions de service public. L'utilisation de ces documents pour d'autres motifs constitue une réutilisation"*) ou encore par le CRPA (*"Les informations publiques figurant dans des documents communiqués ou publiés par les administrations mentionnées au premier alinéa de l'article [L. 300-2](#) peuvent être utilisées par toute personne qui le souhaite à d'autres fins que celles de la mission de service public pour les besoins de laquelle les documents ont été produits ou reçus."*)

La difficulté et l'importance de clarifier la définition de ce terme ont été soulignées dans la littérature, en particulier en ce qui concerne les données de recherche.

Une des questions essentielles porte sur la distinction entre usage et réutilisation. Il existe alors des approches plus restrictives. Certains considèrent par exemple que seuls les usages pérennes de la donnée sont des réutilisations tandis que d'autres considèrent tout usage de la donnée comme une réutilisation.

Dans le cadre de nos missions et de la présente étude, nous considérons comme réutilisation une exploitation des données pour en retirer une information, que cette exploitation soit librement accessible ou non. Ainsi, un usage privé des données ou encore un travail de datajournaliste pour préparer une investigation de terrain sont considérés comme des réutilisations.

Le propos portera ici principalement sur les réutilisations sur [data.gouv.fr](#), c'est-à-dire les réutilisations référencées.

Pourquoi s'intéresser aux réutilisations ?

Les réflexions exposées ici portent sur la connaissance et le référencement de l'existant des réutilisations de données plutôt que sur la question « *comment favoriser les réutilisations ?* », car cette dernière recouvre des sujets bien plus larges (volume des données ouvertes, modalité de mises à disposition, qualité des données, découvrabilité des données, etc.) dont il ne sera pas question ici.

Tout d'abord, la connaissance des réutilisations est essentielle pour mesurer l'impact de l'ouverture des données. Il s'agit de mesurer le nombre de réutilisations et leur qualité, mais aussi leur niveau d'utilisation : s'agit-il d'une application utilisée par des millions de personnes ou d'une visualisation de données ponctuelle par exemple ?

De plus, il existe un besoin récurrent de disposer d'un catalogue de référence de cas d'usage de l'open data. Pour convaincre de la pertinence de la démarche d'ouverture et de son potentiel économique et social, l'ensemble des acteurs ont besoin de s'appuyer sur des exemples concrets.

Le référencement des réutilisations doit également permettre de faciliter la discussion entre réutilisateurs et producteurs de données voire leur collaboration. En effet, les producteurs sont plus disposés à répondre aux réutilisateurs identifiés et dont ils saisissent mieux le besoin. Les réutilisations permettent ainsi d'améliorer la qualité d'un jeu de données en facilitant les remontées de manques et défauts des données.

Enfin, les réutilisations constituent un vecteur d'information au public. Le référencement des réutilisations sur la plateforme nationale permet à un grand nombre d'utilisateurs de découvrir des réutilisations permettant de visualiser les données brutes sous des formes riches et diverses.

PARTIE II : COMPRENDRE LES USAGES ET LES MÉCANISMES DE LEUR RÉFÉRENCEMENT

Notre tour d'horizon des modèles de mesure d'impact de l'open data, combiné à un examen attentif du concept de réutilisation, révèle l'importance de l'analyse des usages dans l'évaluation de l'impact.

Cette deuxième partie dresse une première analyse des usages et de leur possible mesure en s'intéressant aux réutilisations (analyse du catalogue), aux réutilisateurs (identification de communautés) et aux mécanismes qui les encouragent à référencer leurs travaux de réutilisation.

I. ANALYSE DU CATALOGUE DES RÉUTILISATIONS SUR DATA.GOUV.FR

Une analyse du catalogue des réutilisations sur data.gouv.fr a été réalisée durant l'été 2021. Il s'agit d'une matière de travail intéressante bien qu'imparfaite pour analyser les usages de l'open data.

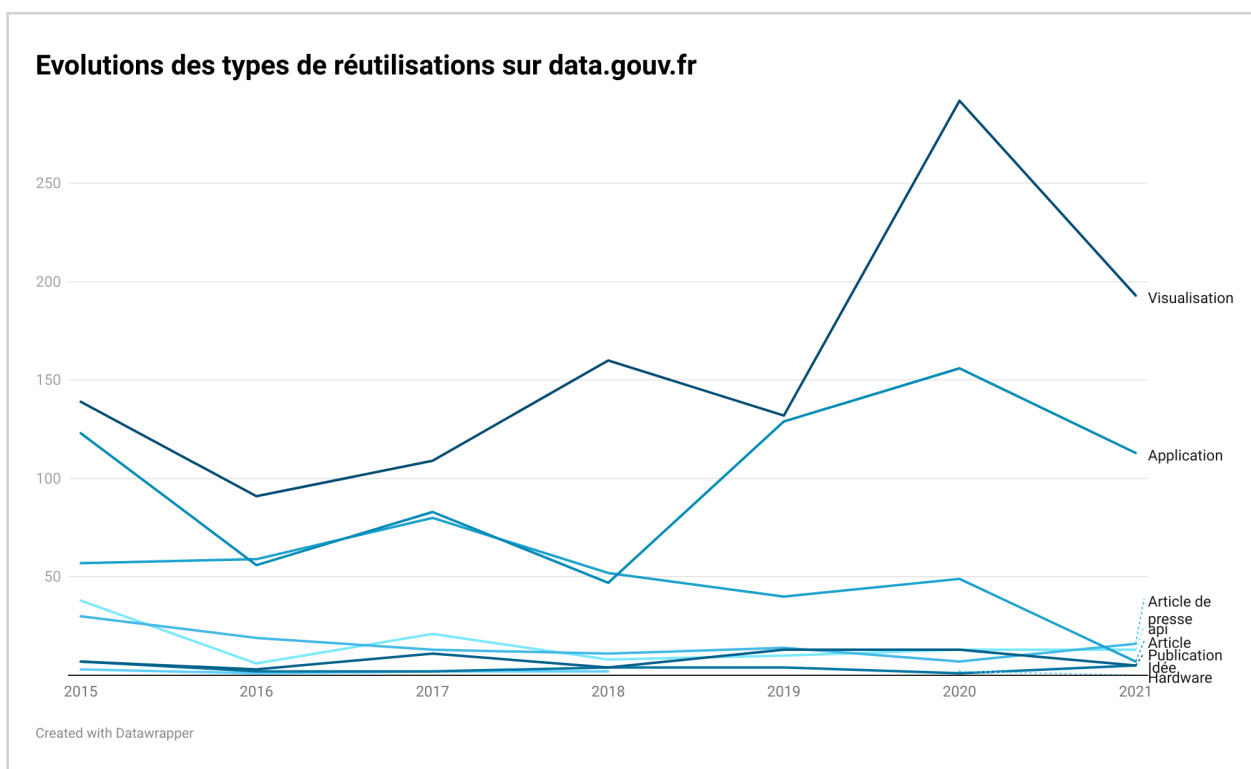
Analyse quantitative

Couverture des usages

Le terme "couverture" désigne ici dans quelle mesure les réutilisations référencées sur data.gouv.fr sont représentatives des usages des données publiques.

Depuis 2013, on observe une moyenne de 332 réutilisations créées par an avec un minimum en 2014 (182 réutilisations). Deux pics peuvent être notés en 2015 et 2020.

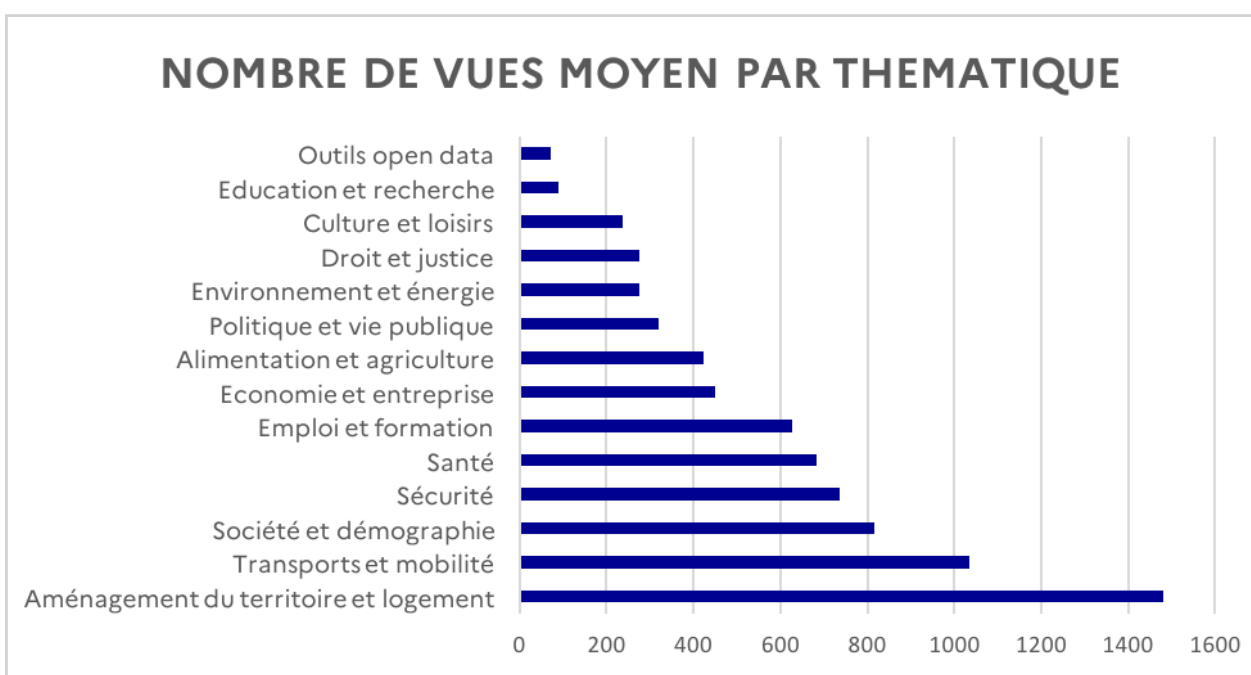
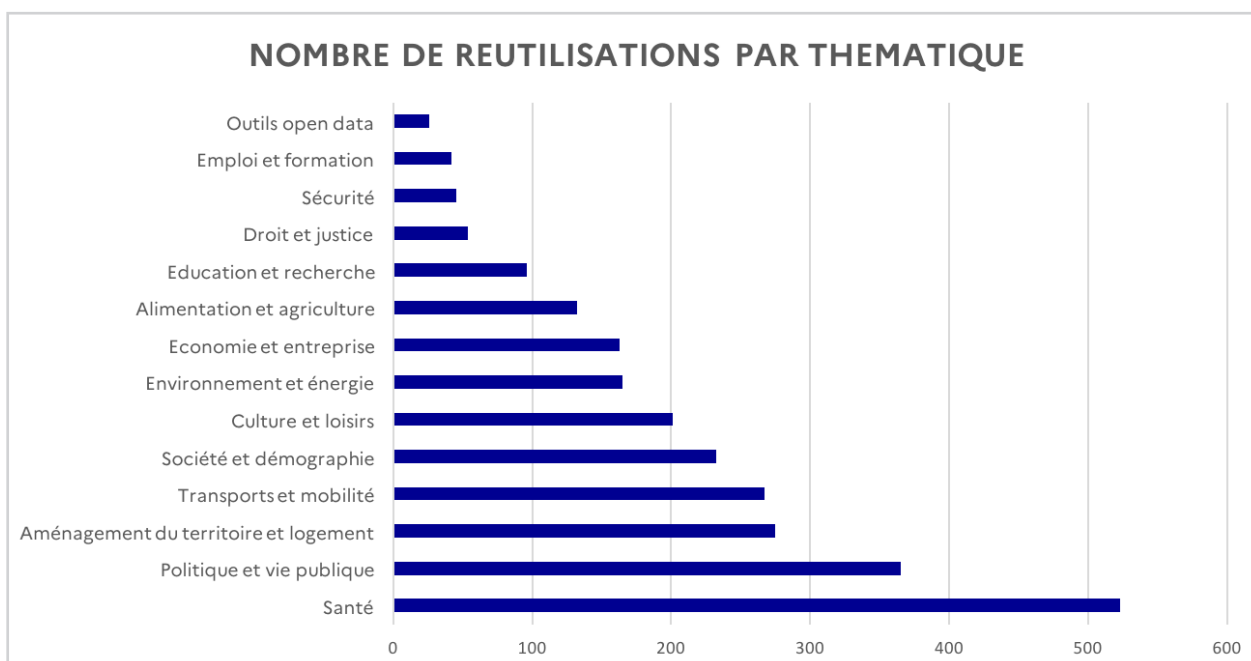
On constate une tendance à la hausse au référencement de réutilisations (environ 600 réutilisations publiées en 2021). Cette augmentation est largement, mais pas seulement tirée par l'engouement autour des données relatives au COVID-19.



Les réutilisations référencées restent peu nombreuses au regard du nombre de jeux de données publiés. Certains jeux de données de qualité et présentant un intérêt certain ne sont pas ou peu accompagnés de réutilisations. On compte en moyenne 0,07 réutilisations référencées par jeu de données.

Certaines thématiques sont bien plus représentées que d'autres.

- Les thématiques recensant le plus grand nombre de réutilisations : Santé (523), Politique et vie publique (365), Aménagement du territoire et logement (275).
- Les thématiques recensant le nombre de réutilisations le plus faible : Droit et justice (54), Sécurité (45), Emploi et formation (42) + Outils open data (26).
- Les thématiques recensant le nombre de vues moyen le plus élevé : Aménagement du territoire et logement (1482), Transports et mobilité (1035), Société et démographie (815).
- Les thématiques recensant le nombre de vues moyen le plus faible : Droit et justice (275), Culture et loisirs (238) et Education et recherche (89) + Outils open data (73).



Typologie des réutilisations

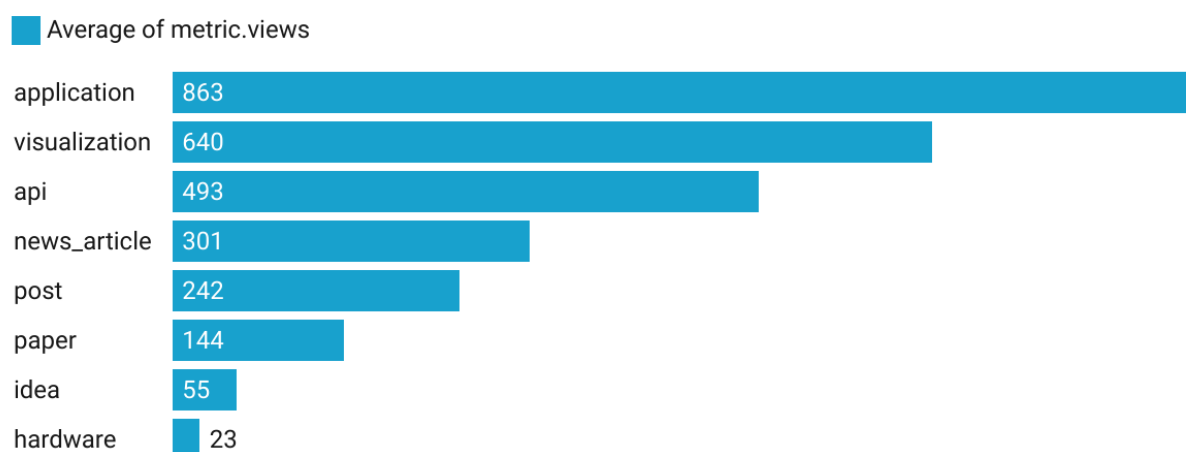
Les visualisations sont le type le plus fréquent (45%), suivies des applications (28%), des articles (13%), puis des API (4%) et des articles de presse (4%), des papiers de recherche (2%), des idées (1%) et enfin du matériel informatique (0,3%).

Les applications et les visualisations rassemblent 88% des vues sur les réutilisations.

Les applications sont moins nombreuses mais plus regardées (différentiel de 11%).

Les articles (posts) sont plus nombreux mais moins regardés (différentiel de 8%).

Visites moyennes par type de réutilisation



Visibilité des réutilisations

On observe sur les réutilisations le même phénomène de longue traîne que sur les jeux de données. La grande majorité des réutilisations n'ont que peu de vues (médiane de 26 vues contre une moyenne à 600), quelques centaines de réutilisations attirent beaucoup de vues (maximum de 112 000).

Il est bon de noter par exemple que 978 réutilisations ont 10 vues ou moins, dont 265 qui n'ont jamais été vues.

Par ailleurs, le nombre de favoris est peu représentatif de l'intérêt actuel pour une réutilisation. Certaines réutilisations anciennes avec de nombreux favoris sont obsolètes. Il s'agit pour la plupart de réutilisations réalisées dans le cadre d'événements comme des hackathons par exemple.

On peut également observer que les réutilisations mises en avant ont en moyenne un nombre de vues 3 fois supérieur.

Visites moyenne selon la mise en avant d'une réutilisation



On constate également une répartition très inégale des discussions sur les réutilisations. Les discussions sont peu nombreuses et concentrées sur quelques jeux de données. Seules 89 réutilisations ont plus d'une discussion.

Analyse qualitative

Les thématiques sont inégalement riches en réutilisations de qualité. La qualité est ici estimée au travers d'une analyse à la main du catalogue ainsi que de l'utilisation de la métadonnée "mise en avant" disponible sur data.gouv.fr. Les thématiques contenant le plus de réutilisations de qualité sont «Santé», «Politique et vie publique», "Aménagement du territoire et logement".

Les thématiques contenant le moins de réutilisations de qualité sont "International", "Education et recherche" et "Culture et loisirs".

Par ailleurs un certain nombre de réutilisations n'ont aucun jeu de données associés (532 réutilisations), souvent car le jeu de données n'existe plus pour des raisons diverses.

Un certain nombre de jeux de données jugés comme importants (soit par notre analyse soit au regard de l'intérêt qu'il suscite estimé par le nombre de vues et de commentaires) n'ont pas ou trop peu de réutilisations de qualité à mettre en avant. On peut citer par exemple les jeux de données suivants :

- [Service-public.fr - Annuaire de l'administration - Base de données locales](#)
- [Transparence-santé](#)
- [Associations Reconnues d'Utilité Publique ARUP](#)
- [Le calendrier scolaire](#)

II. IDENTIFICATION DE COMMUNAUTÉS DE RÉUTILISATEURS

Etablir une typologie de communautés

Les réutilisateurs sont très variés tant en terme d'environnement que de niveau d'expertise et d'usage.

Pour faciliter notre compréhension de leurs usages, en particulier selon le prisme du référencement de réutilisations sur data.gouv.fr, nous avons établi des communautés hybrides selon le facteur identifiant jugé le plus pertinent pour nos activités.

Ce travail d'identification se concentre sur le référencement des réutilisations et ne saurait remplacer le travail plus global de compréhension des usagers de data.gouv.fr, notamment via la mobilisation de concepts issus du design d'expériences utilisateurs (i.e. élaboration de *personas*).

Les communautés suivantes ont été identifiées :

- Les datajournalistes
- Les services publics numériques
- Les enseignants chercheurs
- Les services numériques privés
- Les collectivités territoriales
- Les cabinets de conseil
- Les cartographes
- Les acteurs de la mobilité

Nous avons tenté de décrire ces communautés, en esquisant certaines caractéristiques, leurs intérêts et irritants principaux vis-à-vis des données ouvertes ainsi que leur comportement vis-à-vis du référencement de leurs travaux.

Nous présentons ici deux exemples de communautés identifiées :

Communauté des datajournalistes

Caractéristiques

- La communauté est déjà très structurée, avec un canal Slack et Signal, et une très forte communication entre eux
- Beaucoup d'échanges et de diffusion se font via Twitter
- De plus en plus de journalistes non spécialisés s'intéressent aux données. Leur montée en compétences et l'articulation avec les datajournalistes constitue un enjeu important
- Ils passent beaucoup de temps à nettoyer et récupérer des données (scraping)

Intérêts

- Avoir accès à des données de qualité
- Avoir un contact direct avec le producteur de données
- Avoir les informations rapidement
- Récupérer facilement toutes les mises à jour

Irritants

- Le manque de documentation
- Le manque de fiabilité des données
- Ne pas savoir comment partager des données

Référencement des réutilisations

- Il ne s'agit pas de leur priorité
- L'intérêt leur semble limité si ce n'est pour faire une "bonne action" à destination de la communauté open data
- L'aspect de visibilité potentiel semble dérisoire en particulier pour les grands médias

Communauté des services publics numériques

Caractéristiques

- Faible structuration de la communauté qui s'organise autour de quelques 'pôles' de services publics (comme les Entrepreneurs d'Intérêt Général ou beta.gouv.fr). Certains sont inscrits dans leur administration et sont plus difficilement accessibles.
- Il y a peu d'échanges autour de l'utilisation de données.
- Communauté très diverse, mais aux intérêts et aux problématiques se regroupant par thème.

Intérêts

- API pour alimenter un service de manière simple
- Données mises à jour et fiables
- Connaître les réutilisations qui sont faites des données

Irritants

- Recherche de visibilité
- Manque de temps

Référencement des réutilisations

- Ils connaissent souvent mal les données et la plateforme data.gouv.fr
- L'intérêt à référencer est présent, pour la contribution open data et la visibilité, après un échange en personne.

Mécanismes de référencement des réutilisations

Plus de 40 entretiens spécifiques ont été menés sur la question du référencement des réutilisations.

La trame proposée a été la suivante :

- Questions de profilage
- Questions relatives à la communauté d'appartenance identifiée
- Questions relatives aux usages (type de données, irritants et joies)
- Questions relatives à la relation avec les producteurs de données et leur moyen d'échange
- Questions relatives au référencement sur data.gouv.fr

Bien que la plateforme data.gouv.fr peut s'enorgueillir d'avoir plus de 3 200 réutilisations référencées, les observations issues des analyses et entretiens ont mis en exergue plusieurs problématiques.

Premièrement, la plateforme fait face à une problématique de volume. Les réutilisations référencées ne représentent que le sommet de l'iceberg de l'existant. De nombreux acteurs ignorent encore la possibilité de référencer leur réutilisation ou n'en perçoivent pas l'intérêt.

Deuxièmement, il existe une problématique de qualité des réutilisations. On retrouve ici une certaine similitude avec les problématiques rencontrées sur la qualité des données. À l'instar de la documentation des jeux de données par exemple, de trop nombreuses réutilisations sont accompagnées de descriptions lacunaires qui n'explicitent pas l'usage qui est fait des données.

À ceci s'ajoute une problématique de représentativité des publics parmi les réutilisateurs. Si les acteurs de la société civile sont assez prompts à référencer leurs travaux, d'autres, comme les administrations et le secteur privé par exemple, sont

moins proactifs. Ces derniers représentent pourtant des usages qu'il est essentiel de mesurer.

Par ailleurs, le référencement des réutilisations est complexifié par la diversité des usages. Certains usages étant plus faciles à documenter (cartographie, moteur de recherche) que d'autres (simple observation des données par exemple).

Enfin, le référencement des réutilisations implique également une problématique de modération. *Comment faire la distinction entre publicité et réutilisation ? Comment gérer un contenu changeant dans le temps ?* sont autant de questions qui se posent à l'équipe de la plateforme.

Les entretiens avec les réutilisateurs ont permis de définir les différents avantages du référencement :

- Apporter de l'information au public en orientant les citoyens vers un service s'appuyant sur des données publiques ;
- Donner de la visibilité à la réutilisation et au projet dans lequel elle s'inscrit et démontrer son savoir-faire ;
- Engager un dialogue avec le producteur du jeu de données qui sera plus enclin à répondre aux demandes quand il peut constater de l'usage qui est fait des données.
- Montrer comment le jeu de données peut être réutilisé et inspirer d'autres réutilisations potentielles ;
- Faire avancer l'open data en participant à la construction d'un catalogue de réutilisations permettant d'affirmer l'importance et la pertinence de l'ouverture des données publiques.

PARTIE III : SUIVI ET CONNAISSANCE DES USAGES PAR LES PRODUCTEURS DE DONNÉES

L'étude des usages des données ouvertes nécessite de s'intéresser à l'ensemble des parties prenantes : aussi bien ceux qui font les usages, les réutilisateurs (partie II), que ceux qui les permettent et les favorisent, les producteurs de données.

Le suivi et la connaissance des usages par les producteurs de données constitue donc l'objet de cette troisième partie.

Les producteurs de données suivent-ils les usages de leurs données ouvertes ? Pourquoi s'y intéressent-ils ? Comment les mesurent-ils ? Quels sont leurs besoins en la matière ? Quelles perspectives peut-on tirer de ces enseignements ? Cette partie apporte des éléments de réponse à ces questions.

I. MÉTHODOLOGIE

Les enseignements présentés sont issus d'une étude qualitative qui repose sur des entretiens avec :

- 6 administrateurs ministériels des données, des algorithmes et des codes sources (ministère de l'Europe et des affaires étrangères, ministère de l'agriculture et de l'alimentation, ministère de l'intérieur, ministère de la culture, ministère de la transition écologique, ministère du travail)
- 8 producteurs de données (Base Sirene des entreprises et des établissements, Code officiel géographique, Plan cadastral informatisé, Base adresse nationale, Répertoire opérationnel des métiers et des emplois, Répertoire de l'organisation administrative de l'Etat, Répertoire national des associations)

Les entretiens ont été menés selon la grille d'entretien suivante.

Profil
1. Pourriez-vous me parler de votre poste et de vos principales missions au sein de [nom de l'administration] ?
2. Quelle est votre perception des enjeux de votre administration en termes d'open data et comment y inscrivez-vous vos missions ?
Expérience sur data.gouv.fr
3. A quelle fréquence vous rendez-vous sur data.gouv.fr ?
4. Que venez-vous y faire ?
Usages sur le jeu de données
5. Comment rentrez-vous en contact avec les réutilisateurs du [nom du jeu de données] ? Sur quoi échangez-vous principalement ? Qu'est-ce que cela vous permet de savoir sur vos réutilisateurs et sur ce qu'ils font de vos données ?
6. Pourriez-vous me raconter la meilleure réutilisation du [nom du jeu de données] pour vous ? Et la pire ?

7.	Pourquoi était-ce intéressant pour vous ? Qu'en avez-vous fait et qu'est-ce que cela a-t-il permis de réaliser ?
8.	Comment en avez-vous eu connaissance ?
9.	Comment utilisez-vous la fonctionnalité « Réutilisations » sur data.gouv.fr ?
10.	Si demain vous pouviez accéder de manière exhaustive à tous les usages de vos données, que cherchiez-vous à savoir ? A quoi cela vous servirait-il ?
11.	Si vous pouviez demander tout ce que vous voulez pour mieux connaître les usages de vos données, que demanderiez-vous ?
12.	Avez-vous quelque chose à ajouter ?

II. ETAT DES LIEUX

D'après le rapport « Pour une politique publique de la donnée, des algorithmes et des codes sources », remis au Premier ministre par le député Eric Bothorel, « *les acteurs publics ont aujourd'hui une faible connaissance des réutilisations permises [par les données ouvertes]* » (p.7).

La présente étude décrit et précise les pratiques des producteurs de données quant au suivi et à la connaissance des usages de leurs données ouvertes.

Les données collectées conduisent aux observations suivantes :

- Le niveau de maturité demeure très différent selon le producteur de données : certains disposent d'une bonne connaissance des usages de leurs données (quantitative et qualitative) et se placent comme facilitateurs de nouveaux usages, quand d'autres n'en font qu'une analyse quantitative. Cette hétérogénéité peut résulter d'une différence de sensibilisation aux usages et/ou d'une différence de moyens pour mettre en œuvre le suivi.

- Les producteurs de données portent une attention et mettent en œuvre des efforts croissants pour appréhender les usages. Des producteurs ont par exemple affirmé : « *on aimerait savoir ce que les gens font avec les données et ce dont ont besoin les réutilisateurs* ».
- Les producteurs de données manquent de leviers (méthodologiques, techniques) pour suivre les usages, un constat résumé par un producteur interrogé : « en dehors des métriques de data.gouv.fr, nous avons du mal à observer ce qu'il se passe à l'extérieur de nos murs ».
- Les producteurs de données figurent parmi les premiers réutilisateurs d'autres données ouvertes. Ainsi, les producteurs du Référentiel de l'organisation administrative de l'Etat intègrent sans cesse des données trouvées sur data.gouv.fr dans leur annuaire des services publics et ont mis en place une veille en ce sens.

III. ENJEUX DE LA CONNAISSANCE DES USAGES

Pour les producteurs de données, une analyse des usages permet de développer une offre de données ouvertes plus pertinente, davantage adaptée à la demande. En ce sens, deux mécanismes peuvent être à l'œuvre :

Améliorer la qualité des données

L'observation des utilisations faites et la remontée des besoins des réutilisateurs favorise tout d'abord l'amélioration de la qualité des données. En effet, pour une donnée, la notion de qualité (format, structure, contenu, mise à jour, métadonnées, documentation, etc.) dépend grandement de l'usage qui en est fait. La qualité proposée, bien qu'adaptée aux utilisations internes à l'administration, peut être améliorée pour les usages nouveaux engendrés par l'ouverture. Cette observation est partagée par des producteurs qui ont indiqué : « *nous faisons parfois des choses qui ont du sens pour nous d'un point de vue gestion, mais dont l'utilisation n'est pas optimale* ».

Pôle Emploi travaille étroitement avec la startup d'Etat DiagOriente pour améliorer le Répertoire Opérationnel des Métiers et des Emplois (ROME) en intégrant les retours des utilisateurs de l'outil (compétences pertinentes à retenir, celles qui sont renommées, jamais sélectionnées) et ses travaux de reformulation sémantique des compétences professionnelles.

Motiver la poursuite des efforts d'ouverture

La compréhension des usages permet de motiver la poursuite et le renforcement des efforts de mise à disposition. Les producteurs de données découvrent la valeur de leur travail de publication : les fichiers de gestion de l'administration qu'ils ouvrent deviennent un véritable matériel pour construire des services, des visualisations, des articles journalistiques, etc. Les producteurs de la Base Sirene des entreprises ont expliqué : « *s'il y a des réutilisations qui demandent des téléchargements réguliers, nous pouvons anticiper les conséquences de ce type d'utilisation* ».

A un niveau plus stratégique, les producteurs de données peuvent aussi mobiliser cette connaissance sur les usages pour faire du reporting auprès de leur hiérarchie, démontrer le retour sur investissement des efforts déployés et ainsi justifier la poursuite du travail engagé. Les producteurs de la Base Sirene des entreprises ont précisé valoriser en interne la forte consommation des données, à l'aide d'un bilan mensuel qui reprend le nombre de requêtes et les principaux réutilisateurs.

IV. LEVIERS

MÉTHODOLOGIQUES ET

TECHNIQUES

Les administrations combinent approches quantitatives et qualitatives pour cerner les usages de leurs données ouvertes.

La plupart mesurent les volumes d'usage, en suivant les métriques des jeux de données publiés proposées par data.gouv.fr ou sur leur propre portail : nombre de téléchargements, nombre de réutilisations, etc.

Pour préciser cette analyse, certains producteurs engagent également un dialogue avec les réutilisateurs et investiguent la nature des réutilisations. Selon le niveau de maturité et les moyens disponibles, plusieurs leviers sont mobilisés :

1. Réponses aux commentaires et aux questions soumis sur data.gouv.fr, notamment sur la qualité des données

Il s'agit d'un canal d'échange privilégié par exemple pour les producteurs du Répertoire national des associations, qui compte 344 discussions sur data.gouv.fr.

Suivi des réutilisations ajoutées sur data.gouv.fr et incitation au référencement

La Base adresse nationale recense notamment 43 réutilisations référencées sur la plateforme.

2. Enquêtes auprès des réutilisateurs

Les producteurs du Code officiel géographique ont par exemple lancé en 2019 une enquête qui a permis d'évaluer la satisfaction générale des réutilisateurs et de faire remonter des besoins en termes d'interopérabilité et de fréquence de mise à jour.

Les producteurs de la Base Sirene ont mené en 2022 une démarche similaire et sondé leurs réutilisateurs sur des questions de contenu, de format ou encore de documentation des données mais également sur leur identité.

3. Animation de communautés de réutilisateurs

L'Institut National de l'Information Géographique et Forestière (IGN) organise un certain nombre d'événements mettant à l'honneur les réutilisateurs. Il propose également des conférences, des webinaires de prise en main des différents services ainsi que des tutoriels d'accompagnement.

Pôle Emploi « anime un écosystème large et riche sur l'emploi et le recrutement » (public, privé, recherche, etc.) et dans ce cadre, « est en relation permanente avec [ses] partenaires ».

4. Entretiens avec les principaux réutilisateurs

Les producteurs sondés indiquent manquer de temps pour suivre les usages. Des producteurs ont notamment affirmé : « on aimerait passer encore plus de temps sur l'évaluation de l'usage, mais c'est chronophage ».

Les leviers les plus frugaux, comme les questions soumises sur data.gouv.fr, ne permettent de collecter pour l'instant que peu de retours de la part des réutilisateurs. De plus, selon des producteurs sondés, les questions posées demeurent très spécifiques et ne donnent que peu d'indications sur les travaux réalisés.

V. BESOINS

Des pistes d'amélioration produit ont été remontées pour pallier certaines des problématiques susmentionnées et simplifier le suivi des usages.

Les producteurs de données interrogés plébiscitent ainsi, sur data.gouv.fr :

- Des métriques sur les jeux de données publiés (nombre de consultations, nombre de téléchargements, nombre de réutilisations, etc.)

Plus lisibles et accessibles et communiquées régulièrement sous la forme d'un rapport partageable.

- Des réutilisations

Signalées au moment du référencement (notification) et décrites de façon plus détaillée par les réutilisateurs : utilisation et traitement des données, service rendu, profil du réutilisateur, etc.

- Un forum de discussion qui favorise davantage d'échanges avec les réutilisateurs, et notamment la remontée de leurs besoins.

Les administrations sollicitent également un support méthodologique associé à de l'accompagnement : aide à l'animation de communauté de réutilisateurs, partage de bonnes pratiques sur le suivi des usages, organisation d'échanges entre pairs, etc.

VI. PERSPECTIVES ET OPPORTUNITÉS

Les observations de la présente étude abondent vers une conclusion : les producteurs de données restent les meilleurs ambassadeurs des usages de leurs données ouvertes et les plus à même de suivre ces usages. Il apparaît donc opportun de les autonomiser sur le suivi des usages.

Les pistes suivantes ont été retenues pour l'équipe en charge au sein du département Etalab :

- Sensibiliser les producteurs de données à l'intérêt de connaître les usages en explicitant comment cette connaissance peut être mobilisée ;
- Partager les leviers méthodologiques et techniques ainsi que les bonnes pratiques pour suivre les usages ;
- Proposer un accompagnement sur l'animation de communauté de réutilisateurs ;
- Développer et proposer des outils pratiques pour mesurer les usages, et notamment travailler sur certaines des améliorations produit susmentionnées ;
- Échanger avec les producteurs de données matures sur le sujet, pour collecter davantage de bonnes pratiques.

PANORAMA DES USAGES

Il existe une diversité d'usages et d'impacts potentiels des données ouvertes. Qu'il s'agisse de visualisations, d'applications, d'articles de presse, de papiers de recherche, etc. les réutilisations de données peuvent favoriser la transparence de l'Etat, la transformation de l'action publique, l'information des citoyens ou encore la création de nouveaux services.

Dans le cadre de la présente étude, nous avons entrepris d'approfondir notre connaissance de ces usages en investiguant différents exemples d'utilisation qui illustrent la diversité des sujets traités, des impacts possibles et des personnes qui se saisissent de l'open data.

Leur identification et leur documentation résultent :

- d'un état des lieux des réutilisations sur data.gouv.fr ;
- d'une remontée par des producteurs de données ;
- d'entretiens avec la communauté open data française ;
- d'entretiens avec les réutilisateurs ;
- d'une veille en ligne.

Les cas d'usage retenus sont présentés selon deux prismes d'analyse différents :

1. par secteur
2. par jeu de données à fort impact (service public de la donnée)

I. USAGES SECTORIELS

Usages par l'administration : l'exemple des startups d'État

Les administrations publiques sont les premières utilisatrices des données ouvertes, qu'elles en soient à l'origine ou non.

L'usage au sein de l'administration a ici été analysé à travers l'exemple des startups d'État du programme beta.gouv.fr. Les startups d'État construisent des services publics numériques avec une approche agile qui permet de répondre aux besoins des usagers et de s'adapter en continu.

Leur usage des données ouvertes est pluriel et trois exemples ont été sélectionnés pour l'illustrer.

1. Produire et diffuser des données à forte valeur ajoutée

Certaines startups d'État produisent et exploitent elles-mêmes de nouvelles données. Celles-ci sont mises à la disposition de tous pour permettre leur exploitation mais sont aussi rendues accessibles notamment au moyen de visualisations.

Voici quelques exemples :

- Accès Libre rassemble et cartographie les informations sur l'accessibilité des Établissements recevant du public (ERP) ;
- Camino ouvre les données du domaine minier pour partager l'information sur les projets et faciliter leur gestion ;

- Registre de preuves de covoiturage réunit les données sur le sujet qui sont ensuite notamment exploitées par l'Observatoire du covoiturage ;
- Aides territoires agrège plus de 2 000 aides à destination des collectivités locales et les met à disposition via un moteur de recherche et une API.
- Plus de 18 500 aides sont vues par semaine.

2. Améliorer l'efficacité des services publics

De nombreuses startups d'Etat s'appuient sur des données ouvertes, notamment celles du service public de la donnée pour améliorer l'efficacité des services publics.

La Base Sirene est fortement mobilisée par l'administration pour développer des services nécessitant l'identification facile et rapide des entreprises. Signal Conso permet par exemple aux consommateurs de faire remonter des anomalies. Les informations contenues dans la base Sirene facilitent une identification précise des entreprises concernées. Plus de 100 000 signalements ont été déposés depuis le lancement de la solution.

La Base adresse nationale est également très utilisée par les startups d'Etat, par exemple pour faciliter le remplissage des champs d'adresse grâce à l'autocomplete. Place des entreprises, qui permet aux TPE-PME d'être mises en relation avec le bon conseiller, repose ainsi sur l'API Adresse pour simplifier la recherche de leur SIRET grâce à leur adresse.

3. Développer des services pour protéger notre environnement et notre santé

D'autres startups d'Etat mettent à disposition, exploitent et rendent intelligibles des données sur des thématiques spécifiques. Dans le domaine environnemental, les données brutes deviennent non seulement un véritable moyen d'information sur les enjeux sanitaires et environnementaux mais incitent aussi à une adaptation vertueuse des comportements :

- Recosanté transmet des informations concernant la qualité de l'air sous la forme d'une lettre d'information et des conseils pour se protéger des impacts de la pollution de l'air sur sa santé. Plus de 5 000 abonnés sont aujourd'hui recensés sur Recosanté.
- Datagir utilise divers jeux de données de l'ADEME pour mettre au point des simulateurs qui évaluent l'impact environnemental de nos comportements (alimentation, objets du quotidien, mobilité, télétravail). Plus de 462 000 visiteurs ont consulté les simulateurs depuis le lancement de la solution.

Usages dans le secteur des mobilités et des transports

Dans le domaine des mobilités, les données produites foisonnent. Parmi elles, les données GFTS, pour General Transit Feed Specification (offres de transport en commun au niveau local), et les données de transport en commun en temps réel sont les plus prisées. D'autres données sont aussi très populaires comme les données des infrastructures de recharge pour véhicules électriques ou encore les données concernant l'aménagement de pistes cyclables, en cours d'ouverture.

Deux types d'usage ont été étudiés.

1. Penser les politiques publiques

Au sein de l'administration, les données de transport sont exploitées par des entités spécialisées qui, à partir des données, réalisent des rapports intelligibles et actionnables par les décideurs et le grand public. L'Observatoire national du covoiturage propose par exemple un tableau de bord permettant de suivre, au quotidien, « *l'évolution des pratiques du covoiturage courte distance et d'évaluer l'impact des mesures prises afin d'adapter de façon agile les politiques publiques.* »

Un certain nombre de cabinets de conseil et de bureaux d'études qui accompagnent les collectivités territoriales dans la formulation de politiques publiques d'aménagement mobilisent également ces données. Ils les utilisent pour réaliser leurs travaux de diagnostic : calculer les fréquences de passage à un arrêt de bus, dessiner un plan de réseau, formaliser une structure des flux, etc. Transitec constitue un exemple de ce type d'usage.

2. Alimenter les calculateurs d'itinéraire

La plupart des calculateurs d'itinéraires exploitent des données ouvertes pour concevoir leur service et informer les citoyens : Google Maps, City Mapper, TransitApp, My Bus, Karos, etc. Selon les calculateurs, la couverture territoriale diffère, posant un enjeu d'amélioration de la connaissance des services disponibles dans chaque territoire.

D'autres solutions enrichissent le service avec les données de la base carbone de l'Ademe pour proposer une estimation des émissions carbone de chaque trajet en complément des recommandations d'itinéraire. The Treep peut illustrer un tel usage.

3. Impacts

Les impacts des données de transport sont multiples. Ces données permettent d'accompagner l'évolution des modes de transport et leur démocratisation, à travers trois vecteurs :

- Les données de transport aident à penser et adapter les politiques publiques, notamment sur les thématiques d'aménagement et de mobilité, en interne ou avec l'appui de cabinets d'études.
- Le développement de services numériques s'appuyant sur ces données favorise un accès à une information voyageur plus simple, une mobilité plus multimodale et une couverture spatiale plus égalitaire. L'augmentation de la qualité et de la couverture des données ouvertes permet de réduire les inégalités entre territoires en termes d'accès aux informations de transport. Les données ouvertes ont permis l'accès aux calculateurs d'itinéraire à de nombreuses villes françaises : Figeac, Auxerre, Honfleur, etc.
- L'ouverture de données relatives aux mobilités durables (vélos en libre service, covoiturage) ainsi que le croisement des données de transport avec des données environnementales constituent de puissants leviers d'incitation aux pratiques de mobilité durable.

Si cet impact est par définition difficile à quantifier, l'activité de la plateforme transport. data.gouv.fr peut nous en offrir un premier aperçu : 425 jeux de données recensés et 228 réutilisations référencées. Les chiffres des calculateurs d'itinéraire fournissent également quelques ordres de grandeur : Google Maps revendique 1 milliard d'utilisateurs actifs par mois dans le monde, MyBus en compte 50 000 sur Google Play, City Mapper plus de 10 millions, etc.

De nombreux défis restent à relever pour faciliter l'usage des données de transport : mise à jour des données, découvrabilité, couverture territoriale, etc.

Usages dans l'enseignement supérieur

1. Etat des lieux

De nombreux enseignants s'appuient sur l'open data pour alimenter leurs cours.

En effet, certains domaines d'études requièrent un usage ponctuel ou régulier de données. Les données ouvertes sont notamment une ressource précieuse pour l'analyse d'une politique publique ou d'un phénomène sociologique. On pense par exemple aux données des élections.

Pour d'autres domaines, la donnée et le travail sur la donnée constituent l'objet d'étude lui-même : la géomatique, la data science, etc.

Dans ce cadre, le contenu des enseignements peut se décliner en deux volets :

- Le cadre théorique de l'open data (définition juridique, enjeux, principaux acteurs, plateformes, etc.)
- Le développement de compétences plus techniques liées aux données, qui prend désormais davantage de poids dans les formations. Les enseignants mobilisent plusieurs portails open data, parmi lesquels data.gouv.fr, pour enseigner l'usage et la manipulation de données. Les étudiants apprennent à :
 - Trouver des données publiques pertinentes pour un projet
 - Nettoyer et traiter la donnée
 - Créer des visualisations (graphiques, cartes interactives, etc.)

2. Formats pédagogiques

Il existe de nombreuses manières d'exploiter les données ouvertes dans le cadre de cours. Voici un échantillon de formats pédagogiques :

- Faire rechercher par les étudiants des données pertinentes sur une thématique spécifique (exercice de découvrabilité)
- Proposer des projets autour d'un cas d'usage (sur un semestre entier, avec un rendu final, ou sous forme de séminaire) à l'image des challenges data organisés par Dataactivist
- Proposer aux étudiants de s'entraîner à la préparation et au traitement de données sur des jeux de données qui sont encore difficiles à exploiter à l'état brut. C'est un exercice auquel se sont par exemple attaqués les étudiants de l'université de Grenoble.
- Les données utilisées varient selon le cursus et la spécialité, mais certaines sont particulièrement sollicitées comme les données d'Open Street Map pour la cartographie ou les données locales qui permettent d'étudier différents types d'enjeux à l'échelle d'un territoire (mobilités, inégalités, démographie, etc.)

3. Pistes d'amélioration

L'usage des données ouvertes dans l'enseignement supérieur peut encore davantage se développer.

Si les problématiques de découvrabilité et de qualité peuvent faire l'objet d'un exercice en soi, les efforts visant à les traiter doivent être poursuivis. Les étudiants peuvent y

contribuer en signalant les bugs et en partageant les jeux de données qu'ils ont retravaillé sur data.gouv.fr.

Une meilleure coordination entre enseignants pourrait favoriser la mutualisation des méthodes et des savoir-faire. Le partage des supports pédagogiques entre étudiants et collègues constituerait déjà un premier pas important.

4. Impacts

Mesurer l'impact de l'usage des données ouvertes dans l'enseignement supérieur requiert une méthodologie spécifique, mais voici de premiers axes de réflexion :

- Les étudiants développent des compétences autour des données qu'ils pourront mobiliser dans leur parcours professionnel. A cela s'ajoutent des compétences transverses favorisant l'analyse, l'esprit critique ou la capacité de recherche.
- Ils développent également une connaissance de l'open data et de ses enjeux, contribuant ainsi à son développement et sa compréhension par un plus large public.
- Les travaux réalisés à partir des données permettent de créer de la connaissance, par exemple sur les caractéristiques des villes ou des régions étudiées.

II. USAGES DU SERVICE PUBLIC DE LA DONNÉE

Définition

Le service public des données de référence, créé par l'article 14 de la loi pour une République numérique, vise à mettre à disposition, en vue de faciliter leur réutilisation, les jeux de données de référence qui présentent le plus fort impact économique et social. Ces données sont désormais régies par un régime particulier, afin d'en fiabiliser les modalités d'accès.

La notion de donnée de référence reconnaît le statut particulier de ces bases de données répondant à trois critères : elles servent à identifier ou nommer des produits, des services, des lieux et des personnes ; elles sont utilisées fréquemment par des acteurs publics ou privés autres que l'administration qui les détient ; et la qualité de leur mise à disposition est critique pour ces utilisations.

En 2016, une première liste de neuf jeux de données de référence a été établie : quatre d'entre eux permettent d'identifier des entreprises, des associations, des administrations, des métiers et des emplois, et les cinq autres constituent un corpus relatif aux données géographiques.

Dans le cadre de la présente étude, l'analyse des usages de ces données à fort potentiel d'impact pour l'économie, l'action publique, et la société apparaît particulièrement pertinente. Il s'agit non seulement de mesurer quantitativement l'intérêt pour ces données, mais également de suivre qualitativement les usages.

Etat de lieux des usages sur data.gouv.fr

Il s'agit d'un état des lieux des réutilisations référencées sur data.gouv.fr des neuf jeux de données du service public de la donnée. Celui-ci a été réalisé le 7 mars 2022.

1. Analyse quantitative

L'analyse quantitative est consultable en annexe.

2. Analyse qualitative

Référentiel à grande échelle (RGE)

Ce jeu de données dispose de 4 réutilisations référencées sur data.gouv. Il s'agit de visualisations qui concernent principalement la thématique "Aménagement du territoire et logement". On recense :

- Deux réutilisations intéressantes et de qualité, proposées par le même citoyen réutilisateur (Brice Person). Elles sont du même type et permettent de visualiser l'âge des bâtiments à Paris et dans une petite station balnéaire de Charente Maritime.
 - [Age des bâtiments d'une petite station balnéaire en Charente Maritime](#)
 - [Age des bâtiments de Paris](#)
- Les deux autres ont été réalisées par un agent de l'IGN.

Référentiel de l'organisation administrative de l'Etat

Ce jeu de données dispose de 4 réutilisations référencées sur data.gouv.fr, principalement des applications et des visualisations sur la thématique "politique et vie publique". On recense :

- Une réutilisation de qualité, proposée par les Services du Premier ministre
 - [Annuaire des administrations et des services publics nationaux](#)
- Les autres réutilisations ne sont pas bonnes : deux réutilisations sont inaccessibles et l'autre semble promotionnelle.

Registre parcellaire graphique

Ce jeu de données dispose de 3 réutilisations référencées sur data.gouv.fr sur la thématique "Alimentation et agriculture". On recense :

- Deux réutilisations intéressantes
 - [Ekylibre](#), la solution d'une startup qui propose aux agriculteurs de les aider à gérer leur ferme plus facilement.
 - [Le regroupement des fichiers des différentes éditions du Registre Parcellaire Graphique](#) proposée par un citoyen (Christian Quest) pour simplifier la réutilisation.

- La dernière réutilisation est une visualisation cartographique intéressante du RPG, proposée par l'IGN, mais date de 2013.

Répertoire National des Associations

Ce jeu de données dispose de 11 réutilisations référencées sur data.gouv.fr. Il s'agit d'applications pour moitié et la moitié portent sur la thématique "politique et vie publique". On recense :

- Beaucoup d'annuaires d'entreprises ou plus spécifiquement d'associations, parmi lesquels :
 - [Annuaire entreprises](#) (proposée par Etalab)
 - [Explorateur d'associations](#) (proposée par un citoyen et dont le code est ouvert).
 - [Place des associations](#) (proposée par un blog).
- Des outils intéressants
 - [Une API de recherche dans le RNA](#) (proposée par un citoyen - Christian Quest)
 - [Une API donnée entreprises](#)
- Trois réutilisations ne sont plus accessibles

Plan cadastral informatisé

Ce jeu de données dispose de 11 réutilisations référencées sur data.gouv.fr, toutes sur la thématique "Aménagement du territoire et logement" et majoritairement des applications. On recense :

- Beaucoup de réutilisations qui se rapportent à l'analyse immobilières, combinées aux données des valeurs foncières (beaucoup consultées)
 - [Immo Data](#)
 - [Carte d'exploration des données](#) (proposée par un citoyen - Christian Quest)
 - [Explorateur de biens vendus](#) (proposée par un service privé - Meilleurs agents)
- Des visualisations des parcelles cadastrales (réalisées par Koumoul)
 - [Plan cadastral français](#)
 - [Infos parcelles](#)

Code officiel géographique (COG)

Ce jeu de données dispose de 15 réutilisations référencées sur data.gouv.fr, majoritairement des applications, beaucoup sur la thématique "politique et vie publique". On recense trois réutilisations intéressantes :

- [Annuaire Mairie](#)
- [France découverte](#), l'observatoire cartographique
- [Plateforme de recherche des personnes décédées](#) proposée par un citoyen (Florian Casseyre) dont le code est ouvert

Base Sirene des entreprises et de leurs établissements

Ce jeu de données dispose de 20 réutilisations référencées sur data.gouv.fr, majoritairement des applications et des visualisations sur la thématique "Economie et entreprise".

On identifie un certain nombre de réutilisations intéressantes qui se répartissent ainsi :

- Annales d'entreprises
 - [L'Annuaire des entreprises](#) proposé par la DINUM
- Cartographies de l'activité économique et d'activités en France
 - [Carte interactive des types d'élevage et des régions agricoles](#) proposée par un éditeur de solutions géodécisionnelles (Arcticque)
 - [Carte de la densité des professionnels de santé](#) proposée également par Arcticque
- Services
 - [Signaux Faibles](#), proposé par une startup d'Etat
 - [Place des entreprises](#), proposé par une startup d'Etat

Base adresse nationale (BAN)

Ce jeu de données dispose de 43 réutilisations référencées sur data.gouv.fr, en grande partie des applications, sur des thématiques multiples. On identifie plusieurs réutilisations intéressantes qui se répartissent ainsi :

- Applications de localisation / Calculs d'itinéraire
 - [FindAndGo](#) proposé en open source par une agence de géomatique
 - [Métromobilités](#) proposé par la Métropole Grenobloise
- Des cartographies
 - [panneaux-election.fr](#), proposé par un éditeur de logiciels libre (Mapotempo)
 - [Plaques du matrimoine](#), proposé par un citoyen (Philippe Gambette)
- Des applications d'analyse immobilière
 - [La Loupe Immo](#), service privé

- Autre
 - [In Sun We Trust](#), un calculateur de rendement de panneaux photovoltaïques proposé par un citoyen (Christian Quest)
 - [Open Solar Map](#), proposé par un citoyen (Christian Quest)
- Des outils pour géolocaliser/géocoder des adresses
 - [Plugin expérimental QBano pour QGIS](#), proposée par un citoyen (Christian Quest)
 - [Interface pour géocoder un CSV](#), proposé par Koumoul
- Un outils d'autocomplétion
 - [Plugin d'autocomplétion pour OpenLayers3](#), proposé par Etalab

6 réutilisations ne sont pas accessibles.

Répertoire Opérationnel des Métiers et des Emplois (ROME)

Ce jeu de données dispose de 3 réutilisations référencées sur data.gouv.fr, des visualisations et applications sur la thématique "Emploi et formation".

- Seule une réutilisation est intéressante :
 - [Visualisation des données SKOS du code ROME de Pôle Emploi](#) (proposée par un citoyen - Thomas Francart). La version actualisée 2021 ne fonctionne pas.

Cas d'usage de réutilisation de données ouvertes

Cas d'usage	Réutilisateur	Données utilisées	Problématique et proposition de valeur	Solution	Impact
ERRIAL : Réaliser facilement l'état des risques associé à un bien immobilier	ERRIAL est un service porté par le Ministère de la Transition Ecologique, conçu en partenariat avec le Bureau de recherches géologiques et minières (BRGM).	- Plan Cadastral Informatisé - Données BASIAS - Données sur les sites et sols pollués - Données sur les territoires à risque d'inondation - Données sur les plans de prévention des risques naturels - Données sur les plans de prévention des risques technologiques - Données sur les mouvements de terrain	ERRIAL entend résoudre une problématique prégnante dans le diagnostic immobilier : la dispersion des données relatives aux risques naturels et technologiques. Elle résulte, pour le propriétaire ou le notaire, en une documentation chronophage de l'état des risques obligatoire pour les transactions foncières (ventes, locations). Dès lors, le futur acquéreur ou locataire est souvent informé trop tardivement des risques auxquels est exposé le bien considéré. ERRIAL propose d'exploiter le Plan Cadastral Informatisé pour localiser un bien immobilier, consolider automatiquement les risques naturels et technologiques associés à sa parcelle de terrain et de rendre cette information accessible à tous.	Le Plan Cadastral Informatisé est recoupé avec des données relatives aux risques pour proposer : - Une présentation synthétique et intelligible des risques auxquels un bien immobilier est exposé (sismicité, pollutions du sol, inondations, etc.) - Des cartographies qui positionnent le bien par rapport aux sources de risque - Un pré-remplissage automatique de l'état des risques réglementaire La solution est également proposée sous forme d'API.	Indicateurs d'impact - 320 000 visites/an - 1 000 000 appels à l'API/an Les propriétaires et les notaires peuvent renseigner plus facilement les documents réglementaires relatifs à l'état des risques d'un bien immobilier, avec des informations précises. Les futurs acquéreurs et locataires peuvent rapidement connaître les risques liés aux biens qui les intéressent, en amont de la transmission des documents réglementaires, et prendre des décisions plus éclairées. Tout citoyen peut s'informer et comprendre les risques pouvant affecter son logement.

Nos Villes Vertes : Connaître le patrimoine arboré des communes	'Nos Villes Vertes' a été conçue par Kermap, une startup spécialisée dans la production, l'analyse et la valorisation de données géographiques pour des applications liées à l'aménagement du territoire, l'environnement et l'agriculture.	- Référentiel à grande échelle - Données CORINE Land Cover - Données statistiques sur population française (INSEE)	Nos Villes Vertes entend adresser une problématique croissante : le manque de données exhaustives sur la végétation urbaine, indispensables aux collectivités territoriales pour penser leur politique de végétalisation et informer les citoyens. Nos Villes Vertes propose d'exploiter le RGE pour produire des données complètes sur la végétation arborée de toutes les villes françaises et de rendre cette information accessible à tous.	Les photographies aériennes de la BD Ortho, issue du RGE, ont été analysées avec des modèles algorithmiques qui reposent sur l'apprentissage profond pour contextualiser et extraire automatiquement la végétation arborée. Ces données sont représentées sous la forme de tableaux de bord pour chaque ville française avec - Une cartographie de la végétation arborée - Des indicateurs simples (% de patrimoine arboré, surface par habitant du patrimoine arboré en ville, etc.) - Des visualisations de données	L'exploitation de la BD Ortho permet de produire des données : - objectives, fruit de l'observation factuelle d'une photographie - homogènes, toutes les photographies respectant les mêmes caractéristiques - précises et complètes, couvrant le patrimoine arboré des milieux publics et privés Pour les citoyens : - Une meilleure connaissance de la végétation de leur territoire Pour les collectivités territoriales : - Une vision plus complète de la végétation du territoire - Un premier outil d'aide à la décision pour initier/penser les politiques de végétalisation - La découverte de l'existence de ce type de données et de la possibilité d'y accéder à un niveau adapté à leurs besoins, pour approfondir leurs politiques d'aménagement
---	--	--	---	---	---

La prise d'appel des secours : localiser plus précisément les interventions des pompiers	Le Service départemental d'incendie et de secours des Pyrénées Atlantiques (SDIS64) est en charge de l'analyse des risques et de la mise en place des moyens de secours dans le département.	Base Adresse Nationale	Le SDIS64 s'est saisi d'une problématique critique : le manque de précision relatif à la localisation des interventions des secours. Chaque intervention (secours d'urgence, incendies, etc.) est déclenchée suite à un traitement et une affectation de l'alerte à un Centre d'Incendie et de Secours par le Centre de Traitement de l'Alerte (CTA). Pour chaque appel au 18 qualifié, l'opérateur du CTA localisait une intervention sur une section de voie - appelée tronçon -, en le sélectionnant manuellement dans une liste proposée par le logiciel d'alerte. La plupart des tronçons sont identifiés par un numéro de début et de fin, mais pas tous : il était donc parfois difficile pour l'opérateur de choisir le bon tronçon. Dans les deux cas, les secours étaient envoyés en début de tronçon sélectionné pour chercher le lieu précis sur le terrain, allongeant ainsi leurs délais d'arrivée. Le SDIS64 propose d'utiliser les données adresses issues de la BAN pour localiser les interventions sur le lieu précis et permettre aux secours d'arriver rapidement au bon endroit.	Les données de la BAN ont été intégrées dans le logiciel d'alerte (START) après : - Jointure entre la base des voies et de tronçons (SDIS64) et la Base Adresse Nationale (BAN) - Rapprochement géographique des points adresses de la BAN sur les voies Les opérateurs du CTA peuvent désormais saisir textuellement une adresse dans le logiciel, qui, si elle figure dans la BAN permet : - La sélection automatique du bon tronçon - Le positionnement de l'intervention sur le point adresse, dont les coordonnées sont envoyées aux véhicules d'intervention - Une cartographie pointant sur le lieu exact d'intervention	Indicateur d'impact - 1/2 des interventions du SDIS64 semblent désormais localisées avec le point adresse, d'après l'observation du bulletin quotidien du CTA Pour les pompiers, la localisation des interventions sur le point adresse, plus précise, permet : - L'utilisation de l'itinéraire le plus adapté - La diminution du temps de recherche sur le terrain Et donc la réduction de leurs délais d'arrivée
--	---	-------------------------------	---	--	--

BeeGIS : estimer les ressources florales autour des ruchers	L'Institut technique et scientifique de l'apiculture et de la pollinisation (ITSAP) concourt au développement de la filière apicole à travers l'expérimentation, la recherche appliquée, la formation, etc.	- Registre parcellaire graphique - Registre parcellaire graphique bio - CORINE Land Cover - BD Forêt - BD Topo	BeeGIS entend résoudre un problème récurrent chez les apiculteurs : le manque de connaissance sur les ressources florales à la disposition de leurs abeilles autour des ruchers. Jusqu'alors, ils ne disposaient que d'informations incomplètes, recoupant des investigations réalisées à pied, sur Géoportail ou sur Google Earth. BeeGIS propose de recouper le RPG avec d'autres données ouvertes d'occupation du sol pour une aire de butinage donnée, afin de permettre aux apiculteurs d'analyser simplement les ressources disponibles autour de différents ruchers.	Le RPG constitue l'une des couches géographiques principales d'une application qui propose : - Une cartographie de l'environnement d'un rucher donné (type de sol et plantes cultivées, hydrographie, surfaces bâties, etc.) - Des visualisations de données (surface couverte par type de sol, parts de l'agriculture biologique et du conventionnel par type de culture, etc.)	Indicateurs d'impact - 5 000 apiculteurs professionnels - 1 000 000 ruches La mobilisation de plusieurs jeux de données relatifs à l'occupation du sol, en complément du RPG, permet d'avoir une vision complète de l'environnement d'un rucher : cultures agricoles, zones forestières, zones humides, tissu urbain, etc. L'intégration des données historiques du RPG (2010-2020) permet aux apiculteurs d'anticiper la succession de cultures agricoles associée à une aire de butinage : les évolutions passées présagent les rotations futures. Les apiculteurs : - développent une meilleure connaissance des zones où butinent leurs abeilles - évaluent et comparent les emplacements de rucher à leur disposition pour choisir le plus adapté
---	--	--	--	--	--

Ekylibre : Simplifier la gestion des exploitations agricoles	Ekylibre est une startup spécialisée dans l'édition de logiciels de gestion pour le monde agricole.	- Registre parcellaire graphique - BD Topo - Données E-Phy - Indicateurs de fréquence de traitements phytosanitaires - Données sur le prix des terres	Ekylibre s'attaque à la charge administrative importante que requiert la gestion d'une exploitation agricole. En effet, les agriculteurs doivent aujourd'hui consacrer en moyenne 7 heures par semaine à des tâches administratives (édition de documents réglementaires, saisies de travaux au champ, gestion de trésorerie, etc.). Ekylibre a conçu un outil qui intègre des données ouvertes de référence (RPG, Base Sirene, RGE, etc.) pour automatiser ou faciliter un certain nombre de tâches de gestion.	Une application open source où : - le dossier de l'agriculteur est complété automatiquement (nom, prénom, adresse, etc.) grâce aux données de la Base Sirene - les parcelles de l'exploitation sont cartographiées facilement, en sélectionnant précisément ses zones cultivables à partir du RPG pour offrir : - Tableaux de bord technico-économiques permettant de centraliser, visualiser et analyser l'ensemble des données d'une ferme - Outils interconnectés permettant de gérer entièrement une exploitation agricole : gestion de parcelles, comptabilité, stocks, gestion et traçabilité et la production, etc. - Indicateurs agro-environnementaux et sociétaux	Indicateurs d'impact - 5 000+ agriculteurs - 131 heures par an économisées à la gestion (en moyenne) - 131 742 hectares gérés avec les solutions d'Ekylibre Grâce à Ekylibre et la gestion automatisée des données, les agriculteurs : - voient leur charge administrative réduite - respectent plus facilement les exigences réglementaires - sont en capacité de piloter plus finement, et en temps réel, toutes les activités de leur ferme - peuvent mesurer et améliorer leur impact sur l'environnement
---	--	---	---	--	--

Signaux Faibles : Identifier les entreprises en difficulté	Signaux Faibles est une startup d'Etat portée par un partenariat entre le Ministère de l'Economie, des Finances et de la Relance, le Ministère du Travail, de l'Emploi et de l'Insertion, la Banque de France et URSSAF Caisse Nationale.	- Base Sirene des entreprises et des établissements Les autres données utilisées sont non ouvertes, partagées entre administrations - Données financières, qui apportent une visibilité sur les tendances long terme des entreprises - Données sur l'emploi, qui permettent notamment d'identifier des baisses d'activité - Données sociales, qui peuvent alerter sur des tensions de trésorerie	Signaux Faibles s'attaque à une problématique économique importante : la découverte tardive des difficultés des entreprises. En effet, lorsque celles-ci sont signalées, il est souvent trop tard : beaucoup d'aides publiques qui auraient pu leur permettre de rebondir ne sont plus actionnables ou moins efficaces. Signaux Faibles s'appuie sur l'infrastructure que constitue la Base Sirene pour construire un outil qui permet d'identifier, le plus tôt possible, les entreprises en difficulté. Il devient alors possible de les accompagner précocement et ainsi participer à la prévention des défaillances.	Un algorithme, basé sur des mécanismes d'apprentissage, qui croise et analyse statistiquement des données pour : - repérer et mettre en exergue les signes de fragilité structurelles des entreprises - fournir une prédiction statistique de défaillance à 18 mois et la liste de détection des entreprises en difficulté associée Un ensemble de règles expertes définies avec des utilisateurs et des experts pour capter des signaux ne pouvant être "appris" avec les données historiques. Elles permettent : - d'inclure la détection de signes appréhendés grâce à l'expérience - de favoriser le signalement d'entreprises dont les fragilités correspondent à un symptôme conjoncturel Une application de consultation des données des entreprises et des signalements qui garantit un bon niveau de sécurité dans la diffusion et la consultation des données sensibles Une application de suivi des entreprises/établissements qui facilite la collaboration et l'échange d'informations entre les agents accompagnateurs d'entreprises La Base Sirene permet : - de fiabiliser le référentiel d'entreprises/établissements en associant chaque Sirene/ Siret des jeux de données analysés à une information stabilisée et à jour ; - de compléter le référentiel d'entreprises/établissements apportés par les sources de données analysées avec d'autres champs (Adresse, Code APE, état administratif, etc.) ; - de disposer de l'exhaustivité du stock d'entreprises/établissements, notamment lorsque ces références ne sont pas présentes dans les autres sources de données ; - d'avoir une source fiable sur la localisation des entreprises et ainsi pouvoir affecter les permissions de consultation des informations, qui reposent sur une habilitation territoriale	Indicateurs d'impact - 2 000 000 établissements issus de 500 000 entreprises analysées - 30 000 000 établissements et 20 000 000 entreprises disponibles dans le référentiel - Environ 40% des entreprises suivies par les Commissaires aux Restructurations et à la Prévention des difficultés des entreprises (CRP) ont été identifiées grâce à Signaux Faibles - 500+ agents utilisateurs, dont les 50 CRP La mise en commun de données issues de différentes administrations, complémentaires et adossées à la Base Sirene, permet de disposer de beaucoup d'indices fiables pour détecter les entreprises fragiles. Les agents publics compétents peuvent identifier les entreprises fragiles, investiguer leur situation, cibler leur accompagnement et intervenir au bon moment avec les solutions adaptées (prêt, crédits, accompagnement stratégique, formations, etc.).
--	--	---	---	---	--

DiagOriente : Proposer aux jeunes un parcours d'orientation professionnelle adapté	DiagOriente est une startup d'Etat du Ministère du Travail, de l'Emploi et de l'Insertion, en partenariat avec Id6.	- Répertoire Opérationnel des Métiers et des Emplois - Données sur le marché du travail (besoins en main d'œuvre, nombre d'offres, types de contrat, etc.) - Données sur les offres d'emploi	DiagOriente s'attaque aux problématiques liées à l'orientation scolaire et professionnelle des jeunes, notamment de ceux dits "NEET" (ni en emploi, ni en études, ni en formation). En effet, leur orientation a souvent été "subie" plutôt que "choisie" pouvant conduire à un désengagement, une sortie précoce du système scolaire sans qualification et une entrée sur le marché du travail caractérisée par le chômage et la précarité. Plus généralement, 43% des jeunes n'ont pas de projet professionnel au moment de faire leur choix d'orientation. DiagOriente s'appuie sur le ROME (3.0 & 4.0) comme référentiel des métiers et des compétences pour formuler un outil qui permet aux jeunes de construire activement leur orientation professionnelle en découvrant aptitudes, centres d'intérêt et pistes métiers.	Le ROME (3.0 & 4.0) est distribué dans l'ensemble du parcours de DiagOriente : 1. Identification des compétences - Carte de compétences, à compléter avec des propositions associées aux expériences personnelles, professionnelles et d'engagement consignées 2. Construction du projet professionnel - Sélection et hiérarchisation de centres d'intérêt professionnels - Recommandations par un algorithme de pistes métiers en adéquation avec les centres d'intérêt choisis et les compétences identifiées - Moteur de recherche des métiers associés à des fiches de présentation 3. Recherche/Candidature à des opportunités professionnelles et de formation - Annuaire géolocalisé des emplois, des entreprises accueillant des jeunes en immersion professionnelle et des formations en alternance	Indicateurs d'impact - 63 174 utilisateurs accompagnés - 26 000 cartes de compétences complétées - 532 métiers à découvrir - 72,6% des utilisateurs ont identifié une immersion professionnelle et 30% d'entre eux en ont décrochée une grâce à DiagOriente - 5 800 personnels de l'emploi et de l'orientation dans +326 structures (missions locales, E2C, Epide, etc.) utilisent DiagOriente pour mieux accompagner les jeunes Les jeunes formalisent leurs compétences et intérêts professionnels, découvrent tous les métiers correspondants et trouvent des immersions professionnelles/formations/emplois plus facilement.
--	--	--	---	--	--

EnR de réseaux en Pays de la Loire : Connaître la production d'énergies renouvelables sur son territoire	L'application est le fruit d'un partenariat entre l'Observatoire de la Transition écologique des Pays de la Loire (TEO Pays de la Loire), en charge d'améliorer la connaissance sur l'énergie, le climat, les déchets et les ressources en Pays de la Loire, et la DREAL régionale.	- Registre national des installations de production et de stockage d'électricité - Production électrique annuelle par filière à la maille commune - Points d'injection de Biométhane en France en service - Production annuelle de biométhane par site d'injection (2013-2021) - Production annuelle de biométhane par site raccordé au réseau de GRTgaz et de Teréga - Données locales de consommation d'électricité et de gaz	L'outil s'attaque à une problématique commune aux collectivités territoriales : une appropriation difficile des données relatives aux énergies renouvelables, pourtant essentielles dans une démarche de transition énergétique. Cette difficulté résulte notamment : - d'un accès complexe aux sources disponibles ; - d'une fiabilité parfois relative des données ; - d'un manque d'expertise pour l'exploitation. L'application propose de croiser et de consolider les données produites par les gestionnaires de réseaux d'énergie (Enedis, GRDF, GRTgaz, RTE) pour établir un diagnostic synthétique des énergies renouvelables de réseaux sur le territoire. Elle est déployée à l'échelle de la région des Pays de la Loire.	Un outil open source qui restitue des données relatives aux énergies renouvelables de réseaux préalablement collectées et traitées. Le tableau de bord proposé présente, sur plusieurs attributs du parc de production d'énergies renouvelables de réseaux (nombre d'installations par typologie, puissances installées, consommations couvertes, etc.) : - Des visualisations de données (graphiques, cartes, etc.) - Des indicateurs synthétiques Un service de téléchargement est également disponible pour accéder à l'ensemble des données utilisées (données brutes diffusées par les gestionnaires de réseaux et données traitées).	Pour les collectivités Une meilleure connaissance territoriale des énergies renouvelables qui leur permet d'objectiver les besoins d'évolution de leur mix énergétique et de planifier des actions visant à assurer la transition vers des énergies décarbonées. Pour les gestionnaires de réseaux d'énergie Une valorisation croisée des données qu'ils produisent et diffusent de manière dispersée. Un rapprochement avec les collectivités et autres acteurs territoriaux, qui constituent des partenaires privilégiés dans une transition vers davantage d'énergies renouvelables. Pour tous les citoyens Une connaissance et une transparence accrues concernant les énergies renouvelables produites et distribuées dans leur bassin de vie.
---	--	--	---	---	---

Données de valeur foncière (DVF) : Connaître les prix des biens immobiliers	Etalab est un service du Premier ministre chargé notamment de coordonner la conception et la mise en œuvre de la stratégie de l'Etat dans le domaine de la donnée.	- Demandes de valeurs foncières - Plan Cadastral Informatisé	L'outil répond à une problématique identifiée lors de l'ouverture de la base « Demande de valeurs foncières » : les données, diffusées en open data pour favoriser la transparence des marchés fonciers et immobiliers, restaient difficilement manipulables et accessibles par le grand public. Or, 72% des français souhaitent connaître les prix de vente des biens de leur quartier. L'outil propose de cartographier la base « Demandes de valeurs foncières » à l'aide du Plan Cadastral Informatisé pour permettre de trouver facilement des informations claires sur les biens vendus et ainsi les rendre accessibles à tous.	Une carte interactive qui affiche : - La liste des transactions immobilières et foncières associées à une zone – parcelle cadastrale – donnée : appartements, maisons, parcelles, exploitations, etc. - Des informations détaillées pour chaque transaction : prix de vente, date de la transaction, adresse, descriptif du bien Le code de l'application est en open source.	Les acheteurs et vendeurs peuvent mieux évaluer la valeur d'un bien immobilier, en se rapportant à la valeur foncière de biens similaires, et ainsi prendre des décisions plus éclairées. Les notaires et agences immobilières peuvent mieux conseiller leurs clients et notamment fixer des prix de vente plus justes et cohérents. Tout citoyen peut s'informer sur le prix de vente d'un logement ou d'un terrain.
---	---	---	---	--	---

Annuaire des Entreprises : Retrouver facilement les informations sur les entreprises	Etalab est un service du Premier ministre chargé notamment de coordonner la conception et la mise en œuvre de la stratégie de l'Etat dans le domaine de la donnée.	- Base Sirene des entreprises et de leurs établissements - Base Adresse Nationale - Répertoire national des associations - Registre National du Commerce et des Sociétés - Répertoire National des Métiers - Données BODACC - Annonces et comptes annuels des associations - Comptes associations - Conventions collectives	L'Annuaire des Entreprises s'attaque à une problématique partagée par tous (agents publics, entrepreneurs, particuliers, etc.) : la dispersion des données relatives aux entreprises, qui rend difficile et chronophage la recherche d'information. L'outil propose de centraliser toutes les données ouvertes relatives aux personnes morales qui figurent dans les données de l'administration (entreprises, associations, administrations, entrepreneurs, etc.) et de mettre ces informations consolidées à la disposition de tous, gratuitement.	L'ensemble des données sont récupérées par API pour alimenter un site public qui propose : - Un moteur de recherche performant pour retrouver rapidement une personne morale - Une page dédiée à chaque personne morale, avec toutes les informations légales la concernant (SIREN/SIRET, n° de TVA, code APE, justificatif d'immatriculation, dirigeants, conventions collectives, etc.)	Les données en open data permettent de disposer de sources officielles, fiables et exhaustives, s'agissant des informations détenues par l'administration. Les agents publics qui ont besoin d'informations sur les personnes morales dans le cadre de leurs missions peuvent travailler plus efficacement (vérifier l'existence d'une personne morale, certifier sa validité, etc. dans le cadre des marchés publics, de la lutte contre la fraude, etc.). Les entrepreneurs peuvent récupérer leurs informations rapidement et ainsi effectuer leurs démarches administratives plus facilement (SIREN/SIRET, code NAF/APE, n° de TVA, etc.). Les citoyennes et citoyens peuvent s'informer et vérifier les informations légales des personnes morales avec lesquelles ils et elles sont en relation.
--	---	---	--	---	--

CONCLUSION

La compréhension des usages des réutilisateurs de données, le référencement et la valorisation de ces usages sont des missions essentielles pour plusieurs raisons.

Cela permet notamment de mobiliser les producteurs dans l'ouverture des données, mais aussi de mesurer l'impact des efforts d'ouverture.

Les efforts d'accompagnement sur le sujet des réutilisations à destination des producteurs et des réutilisateurs de données doivent avancer de façon parallèle et cohérente aux évolutions de la plateforme data.gouv.fr pour favoriser le référencement et mieux mesurer les usages.

Accompagner les producteurs et réutilisateurs de données

Accompagner les producteurs de données

Au même titre qu'Etalab accompagne les producteurs de données dans l'ouverture de leurs données, il se doit de poursuivre l'accompagnement dans le cycle de l'appropriation de ces données par des parties prenantes tierces.

L'objectif doit être d'autonomiser les producteurs de données dans la valorisation de leurs données et dans l'animation de la communauté des réutilisateurs.

Ceci passe notamment par développer la culture de la réutilisation des données ouvertes : comprendre et faire comprendre les usages

Ceci nous a notamment poussé à proposer un catalogue de cas d'usages, déclinés par politique publique, évoluant dans le temps, pour incarner les réutilisations. Un effort particulier a été fait sur la compréhension des usages des jeux de données de qualité, au premier rang desquels les jeux de données du Service Public de la Donnée.

L'objectif doit être
d'autonomiser
les producteurs
de données dans
la valorisation de
leurs données et
dans l'animation
de la communauté
des réutilisateurs.

Il s'agit également de **mettre à disposition une offre de service à destination des producteurs sur la valorisation de leurs données et l'animation de communauté.**

Il s'agit d'étendre la notion de data literacy à la valorisation des données et à l'animation de communauté. Cette offre de service s'appuiera notamment sur un accompagnement sur la sélection des données à valoriser, la mise en contact avec les communautés pertinentes, le partage de bonnes pratiques, l'organisation d'événements, etc.

Le succès de cet objectif peut être mesuré de plusieurs façon : le nombre de producteurs engagé dans une démarche proactive de valorisation des données et de compréhension des usages; Le nombre de producteurs accompagné dans la valorisation de leurs données; La qualité du catalogue de cas d'usages variés et actionnables; La satisfaction des producteurs de données vis-à-vis de l'offre de service.

Accompagnement des réutilisateurs de données

Si la dimension d'animation de communautés doit avant tout passer par le biais des producteurs de données, il s'agit d'être en capacité de mobiliser les réutilisateurs de manière proactive.

Plusieurs actions ont été identifiées pour atteindre cet objectif.

Parmi celles-ci de **communiquer sur l'aspect communautaire de data.gouv.fr et l'intérêt de référencer des réutilisations.** Ce positionnement doit être clair sur tous les médias où data.gouv.fr communique et dans les interactions avec nos usagers.

Il s'agit également de formaliser une stratégie d'animation et de communication pour susciter l'intérêt des communautés de réutilisateurs. Il faudra donc définir les communautés à cibler et les canaux privilégiés afin de favoriser par exemple l'organisation d'événements avec les producteurs, de formations ou encore l'identification de réutilisateurs « ambassadeurs ».

Enfin il nous faut constamment **mieux accompagner les réutilisateurs et répondre à leurs demandes.** Même si l'accompagnement sera aussi amélioré par

les adaptations produites, il est nécessaire d'adapter le support aux demandes et d'enrichir la documentation pour mieux répondre aux sollicitations des réutilisateurs.

Le succès de cet objectif peut être mesuré de plusieurs façon : le nombre d'événements organisés, le délai de réponse au support, le nombre de réutilisations référencées.

**L'objectif est
d'animer la
communauté des
réutilisateurs
de données
afin de faciliter
l'appropriation
des données et
de gagner en
connaissances sur
les usages.**

Améliorer l'expérience des réutilisations de données sur data.gouv.fr

Améliorer la découvrabilité des réutilisations

Il existe un problème de couverture du catalogue : trop peu de réutilisations sont référencées sur data.gouv.fr et celles-ci sont souvent similaires (en termes de type, de thème et d'auteur). Par exemple, les administrations et le secteur privé sont peu représentés. D'autre part, les réutilisations sont trop peu découvrables sur la plateforme.

Ceci passe par rendre les réutilisations plus visibles dans la navigation notamment au niveau de chaque jeu de données et par producteur.

Mais aussi par une **refonte du système de vote communautaire sur la plateforme** pour répondre à des problématiques de découvrabilité, mais aussi inciter au référencement.

Il s'agit également de **développer l'éditorialisation et la mise en récit des meilleures utilisations de jeux de données** en adoptant une approche hybride entre réutilisations et mise en avant de cas d'usages.

Ces actions peuvent notamment être mesurées par les évolutions des visites sur les réutilisations ainsi que les évolutions des réutilisations référencées.

Améliorer la qualité des réutilisations

Il y a des déficits de qualité des réutilisations. Une partie très importante des réutilisations référencées sont du one shot, les applications par exemple ne sont que rarement maintenues, les descriptions sont souvent lacunaires et l'usage fait des données est rarement explicité. D'ailleurs, la grande majorité des réutilisations n'est que rarement consultée.

L'objectif ici est d'améliorer la diversité et la qualité des réutilisations référencées sur la plateforme.

Pour ce faire, il s'agit de **repenser le parcours de publication des réutilisations** notamment mieux accompagner les réutilisateurs dans la documentation de leurs travaux, enrichir les formats de description, etc.

Il faut également enrichir la documentation existante et fournir des exemples. À ce titre les cas d'usages peuvent être précieux. L'amélioration de la qualité des réutilisations passe autant par l'accompagnement de ces derniers que par des évolutions produits.

Un moyen de mesurer cet objectif est de suivre le nombre de jeux de données du top 50 accompagné d'une réutilisation de qualité.

**Il nous faut
faciliter la
découverte et la
navigation des
réutilisations de
données.**

Favoriser les échanges entre producteurs et réutilisateurs de données

data.gouv.fr est un espace particulièrement opportun pour être une plateforme d'échanges entre producteurs et réutilisateurs de données. Ces échanges s'inscrivent dans un cercle vertueux qui permet d'améliorer la qualité des données.

Nous visons notamment **d'améliorer le système de commentaires de data.gouv.fr**, notamment, permettre les votes, les mentions, des fonctionnalités de modération, etc.

Il peut également être opportun d'explorer la possibilité de **permettre aux réutilisateurs de donner une indication sur l'usage pressenti des données**.

Une des mesure d'impact ici peut être l'évolution du nombre de commentaires sans réponses.

L'objectif est que
les producteurs
répondent
davantage aux
questions et que
leur expérience
soit améliorée
comme celle des
réutilisateurs.

BIBLIOGRAPHIE SÉLECTIVE

Cette bibliographie n'est pas exhaustive.

Études et rapports

The Research Institute of the Finnish Economy (2015), The impact of open data –a preliminary study, https://www.w3.org/2013/share-psi/wiki/images/6/67/Impact_of_Open_Data_in_the_Public_Sector_Koski_2015.pdf

Open data barometer, Leaders Edition (2018) <https://opendatabarometer.org/doc/leadersEdition/ODB-leadersEdition-Report.pdf>

Open Data Barometer, Leaders Edition - Research Handbook (2018) https://docs.google.com/document/d/1I24b20z9n3-FF8UCZjqv3pB6_ENmJYROZ_0cyQIAWM/

The Economic Benefits of Open Data (2017) https://www.europeandataportal.eu/sites/default/files/analytical_report_n9_economic_benefits_of_open_data.pdf

OECD report : Open, Useful and Re-used data (OURdata) Index (2019) <https://www.oecd-ilibrary.org/docserver/45f6de2d-en.pdf?expires=1588285270&id=id&accname=guest&checksum=4BECFDAB1DD6A1A64365369310DC7E5B>

Chapter 8 of OCDE report (2019) <https://www.oecd-ilibrary.org/docserver/9789264305847-11-en.pdf?expires=1588285747&id=id&accname=guest&checksum=8B3BA94C785AD6B1BAFD6875D855FB1F>

Open Data Impact Map (2016) <http://reports.opendataenterprise.org/May2016Report.pdf>

Fostering a Critical Development Perspective on Open Government Data (2012) <https://idl-bnc-idrc.dspacedirect.org/bitstream/handle/10625/56317/IDL-56317.pdf>

Open Data Institute - case studies <https://theodi.org/knowledge-opinion/case-studies/>

Dix ans d'ouverture des données publiques | Fing <https://fing.org/toutes-les-actions/open-data-impact.htm>

OCDE. Rapport sur les données ouvertes publiques Encourager la maturité des politiques de données ouvertes pour un impact durable. Paris : OECD Publishing. 2019. <https://www.oecd.org/fr/innovation/rapport-sur-les-donnees-ouvertes-publiques-12ea5027-fr.htm>

A Guide to Open Government and the Coronavirus: Open Data. (2020, July 01). Retrieved from <https://www.opengovpartnership.org/documents/a-guide-to-open-government-and-the-coronavirus-open-data/>

In times like these, transparency matters more than ever - News. (n.d.). Retrieved from https://www.transparency.org/news/feature/in_times_like_these_transparency_matters_more_than_ever?utm_medium=email&utm_campaign=The

Knowledge & opinion. (n.d.). Retrieved from <https://theodi.org/knowledge-opinion/case-studies/>

Open data barometer, Leaders Edition (2018). <https://opendatabarometer.org/doc/leadersEdition/ODB-leadersEdition-Report.pdf>

Open Data Barometer, Leaders Edition - Research Handbook (2018). Retrieved from : https://docs.google.com/document/d/1I24b20z9n3-FF8UCZjzqv3pB6_ENmJYROZ_0cyQIAWM/

The Economic Benefits of Open Data (2017) Retrieved from : https://www.european-dataportal.eu/sites/default/files/analytical_report_n9_economic_benefits_of_open_data.pdf

OECD report : Open, Useful and Re-usable data (OURdata) Index (2019). <https://www.oecd-ilibrary.org/docserver/45f6de2d-en.pdf?expires=1588285270&id=id&accname=guest&checksum=4BECFDAB1DD6A1A64365369310DC7E5B>

Chapter 8 of OCDE report (2019) .<https://www.oecd-ilibrary.org/docserver/9789264305847-11-en.pdf?expires=1588285747&id=id&accname=guest&checksum=8B3BA94C785AD6B1BAFD6875D855FB1F>

Open Data Impact Map (2016). <http://reports.opendataenterprise.org/May2016Report.pdf>

Fostering a Critical Development Perspective on Open Government Data (2012) <https://idl-bnc-idrc.dspacedirect.org/bitstream/handle/10625/56317/IDL-56317.pdf>

OECD (2019b). The Path to Becoming a Data-Driven Public Sector, OECD Digital Government Studies, OECD Publishing, Paris. <https://doi.org/10.1787/059814a7-en>

OECD (2020), Open, Useful and Re-usable data (OURdata) Index: 2019, OECD Public Governance Policy Papers, No. 01, OECD Publishing, Paris. <https://doi.org/10.1787/45f6de2d-en>

International Open Data Charter (2015). Principles, International Open Data Charter, viewed on 15 July 2020. <https://opendatacharter.net/principles>

Ubaldi, B. (2013), «Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives», OECD Working Papers on Public Governance, No. 22, OECD Publishing, Paris. <https://doi.org/10.1787/5k46bj4f03s7-en>

Eaves, D., & McGuire, B. (2019). The first decade of open data has been a win — but not for the reasons you think, [Apolitical.co](https://apolitical.co), viewed on 15 July 2020, <https://apolitical.co/en/solution_article/the-first-decade-of-opendata-has-been-a-win-but-not-for-the-reasons-you-think

van Ooijen, C., B. Ubaldi and B. Welby (2019). A data-driven public sector: Enabling the strategic use of data for productive, inclusive and trustworthy governance, OECD Working Papers on Public Governance, No. 33, OECD Publishing, Paris. <https://doi.org/10.1787/09ab162c-en>

D'Auria, G., & De Smet, A. (2020). Leadership in a crisis: Responding to the coronavirus outbreak and future challenges, McKinsey & Company, viewed on 16 July 2020. <https://www.mckinsey.com/businessfunctions/organization/our-insights/leadership-in-a-crisisresponding-to-the-coronavirus-outbreak-and-future-challenges#>

Zeenea (n.d). How does data democracy strengthen agile data governance?, Zeenea White Paper on Agile Data Governance, Volume 2 : Agile Governance Roles and data cultures. <https://zeenea.com/project/how-does-data-democracy-strengthen-agile-data-governance/>

Data Europa. Open Data Maturity report 2021. 2022. https://data.europa.eu/sites/default/files/landscaping_insight_report_n7_2021_0.pdf

Articles scientifiques

Rivera Perez, J. (2015), Beyond Open Data Disclosure: Fostering The Impact Of Open Government Data Towards More Efficient Public Institutions, London School of Economics and Political Science, Department of Management, London.

Zuiderwijk, A. & Janssen, M. (2014), Open data policies, their implementation and impact: A framework for comparison, Government Information Quarterly, Volume 31, Issue 1, Pages 17-29URL : <https://www.sciencedirect.com/science/article/abs/pii/S0740624X13001202>

Pierre-Henri Morand. La dématérialisation des marchés publics : objectif(s) atteint(s) ?. État, droit et mutation numérique : regards croisés sur la dématérialisation, PUAM, inPress. <https://hal.archives-ouvertes.fr/hal-03392175/>

Flavien Lucas, Michell Queiroz, Kenneth Sörensen. ReQreate : utilisation de données ouvertes pour générer des instances réalistes de transport à la demande.. 23ème congrès annuel de la Société Française de Recherche Opérationnelle et d'Aide à la Décision, INSA Lyon, Feb 2022, Villeurbanne - Lyon, France. Ffhal-03595401f <https://hal.archives-ouvertes.fr/hal-03595401/document>

Arthur Sarazin. Construire les usages des données ouvertes en France : proposition d'un Open Data Canvas par l'approche du design science. Gestion et management. Université Grenoble Alpes [2020-..], 2021. Français. ffNNT : 2021GRALG007ff. Fftel-03496226f <https://tel.archives-ouvertes.fr/tel-03496226/document>

Les effets de l'open data et du Big Data dans la fabrique des cartes de l'action publique : entre repositionnement des acteurs et démonopolisation de l'expertise autour des

données territoriales; Boris Mericskay <https://halshs.archives-ouvertes.fr/halshs-03464763/document>

Saint Auguste, S. (2021). L'enjeu de l'*open data* pour les éditeurs et les *legaltech* : une redéfinition du marché de l'information juridique. *Légipresse*, 65, 87-93. <https://doi.org/10.3917/legip.hs65.0087>

Jensen, Michael C et William H Meckling. « Theory of the firm: Managerial behavior, agency costs and ownership structure », *Journal of financial economics*. 1976, vol.3 no 4. p. 305- 360. <https://www.sciencedirect.com/science/article/pii/0304405X7690026X>

Chevallier, Jacques. « Vers l'État-plateforme ? », *Revue française d'administration publique*. 2018, vol.167 no 3. p. 627- 637. https://catalogue-bibliotheque.sciencespo.fr/permalink/33USPC_SPO/1gsggji/cdi_crossref_primary_10_3917_rfap_167_0627

Algan, Y., Bacache-Beauvallet, M. & Perrot, A. (2016). Administration numérique. Notes du conseil d'analyse économique, 34, 1-12. <https://doi.org/10.3917/ncae.034.0001>

L'open data de crise : entre mobilisation citoyenne et communication gouvernementale; Simon Chignard <https://www.annales.org/enjeux-numeriques/2021/en-2021-06/2021-06-11>.

HOANG, M., & GARRONE, A. Le service public des données de référence: enjeux de gouvernance. DES MINES, 39. https://www.annales.org/enjeux-numeriques/2022/en_18_06_22.html

Rivera Perez, J. (2015), Beyond Open Data Disclosure: Fostering The Impact Of Open Government Data Towards More Efficient Public Institutions, London School of Economics and Political Science, Department of Management, London.

Zuiderwijk, A. & Janssen, M. (2014), Open data policies, their implementation and impact: A framework for comparison, *Government Information Quarterly*, Volume 31, Issue 1, Pages 17-29.

URL : <https://www.sciencedirect.com/science/article/abs/pii/S0740624X13001202>

Pasquetto, I. V., Randles, B. M., & Borgman, C. L. (2017). On the Reuse of Scientific Data. *Data Science Journal*, 16, 8. DOI: <http://doi.org/10.5334/dsj-2017-008>

Van de Sandt, S., Dallmeier-Tiessen, S., Lavasa, A., & Petras, V. (2019). The Definition of Reuse. *Data Science Journal*, 18(1), 22. DOI: <http://doi.org/10.5334/dsj-2019-022>

Ouvrages fondamentaux

Tim O'Reilly, « Government as a Platform », Innovations: Technology, Governance, Globalization, vol. 6, no 1, 2011, p. 13-40 (DOI 10.1162/INOV_a_00056)

Verdier Henri. La donnée comme infrastructure essentielle : Rapport au premier ministre sur la donnée dans les administrations 2016-2017. 2018. https://www.etalab.gouv.fr/wp-content/uploads/2018/04/RapportAGD_2016-2017_web.pdf

Autres articles et ressources

Pichot Damon, E. (2019), "Des analytics à l'impact, pourquoi évaluer une politique open data ?", article de blog, URL : <https://medium.com/dataactivist/des-analytics-à-l-impact-pourquoi-évaluer-une-politique-open-data-461504a34eca>

Gaylord Marchesseau. « Opinion | Rapport Bothorel : quelles suites pour l'open data ? », Le Cercle Les Echos. 2021. <https://www.lesechos.fr/idees-debats/cercle/opinion-rapport-bothorel-queelles-suites-pour-l-open-data-1294759>

La réutilisation des données. (n.d.). Retrieved from [La réutilisation des données](#)

Quand la donnée arrive en ville. (n.d.). Retrieved September 26, 2022, from https://books.google.fr/books?hl=fr&lr=lang_fr&id=EKoWEAAQBAJ&oi=fnd&pg=PT3&dq=open+data&ots=MPdMG7PulA&sig=F8IST1Gors-Ap_kXzkaqNO_Qy8&redir_esc=y#v=onepage&q=open

Laboratoire d'Analyse et de Décryptage du Numérique | Mission Société Numérique. (2020, March 30). Ouverture des données publiques relatives au Covid19 : Les premières réutilisations voient le jour - Laboratoire d'Analyse et de Décryptage du Numérique : Mission Société Numérique. Retrieved from <https://labo.societenumerique.gouv.fr/2020/03/30/des-donnees-relatives-au-covid19-en-open-data-les-premieres-reutilisations-voient-le-jour/>

ANNEXES

Analyse des réutilisations sur data.gouv.fr du Service Public de la Donnée

Nom du jeu de données	Producteur	Date de création	Nombre de vues	Nombre de discussions	Nombre de réutilisations	Type de réutilisations	Thématiques des réutilisations	Nombre de vues des réutilisations
Base Adresse Nationale	Services de la Première ministre	17/04/2015	23171 (Moyenne = 125 ; Médiane = 0)	65	43	- Application (28) - Visualisation (8) - API (6) - Post (1)	- Outils open data (9) - Aménagement du territoire (7) - Politique et vie publique (7) - Santé (5) - Economie et entreprise (5) - Société et démographie (3) - Transports et mobilité (2) - Environnement et énergie (2) - Culture et loisirs (2) - Droit et justice (1)	- Moyenne : 2396 - Médiane : 108 - Min : 0 - Max : 68577
Base Sirene des entreprises et de leurs établissements	Institut National de la Statistique et des Etudes Economiques (INSEE)	24/08/2018	28213 (Moyenne = 125 ; Médiane = 0)	15	20	- Application (10) - Visualisation (7) - API (2) - Post (1)	- Economie et entreprise (13) - Alimentation et agriculture (2) - Santé (2) - Politique et vie publique (2) - Outil open data (1)	- Moyenne : 365 - Médiane : 34 - Min : 0 - Max : 2027
Code Officiel Géographique	Institut National de la Statistique et des Etudes Economiques (INSEE)	15/03/2017	13405 (Moyenne = 125 ; Médiane = 0)	13	15	- Applications (9) - Visualisations (4) - API (2)	- Aménagement du territoire (3) - Culture et loisirs (1) - Outils open data (1) - Politique et vie publique (6) - Santé (1) - Société et démographie (2) - Transports et mobilité (1)	- Moyenne : 1458 - Médiane : 242 - Min : 0 - Max : 8838
Plan Cadastral Informatisé	Ministère de l'économie des finances et de la relance	06/04/2017	72843 (Moyenne = 125 ; Médiane = 0)	23	11	- API (2) - Applications (7) - Visualisations (2)	Aménagement du territoire	- Moyenne : 9466 - Médiane : 3913 - Min : 66 - Max : 68577

Registre parcellaire graphique	Institut National de l'Information Géographique et Forestière	27/03/2017	23422 (Moyenne = 125 ; Médiane = 0)	33	3	- Application (1) - Post (1) - Visualisation (1)	Alimentation et agriculture	- Moyenne : 862 - Médiane : 270 - Min : 89 - Max : 2226
Référentiel de l'organisation administrative de l'Etat	Services de la Première Ministre	12/05/2016	1711 (Moyenne = 125 ; Médiane = 0)	10	4	- Application (2) - Visualisation (2)	- Environnement et énergie (1) - Politique et vie publique (3)	- Moyenne : 28 - Médiane : 30 - Min : 0 - Max : 53
Référentiel à grande échelle	Institut National de l'Information Géographique et Forestière	06/04/2017	1123 (Moyenne = 125 ; Médiane = 0)	1	4	Visualisations	- Aménagement du territoire (3) - Politique et vie publique (1)	- Moyenne : 147 - Médiane : 146 - Min : 0 - Max : 294
Répertoire national des associations	Ministère de l'Intérieur	05/04/2017	20489 (Moyenne = 125 ; Médiane = 0)	331	11	- API (3) - Application (5) - Visualisation (3)	- Economie et entreprise (2) - Education et recherche (1) - Politique et vie publique (5) - Société et démographie (2) - Transports et mobilité (1)	- Moyenne : 663 - Médiane : 24 - Min : 0 - Max : 6346
Répertoire opérationnel des métiers et des emplois	Pôle Emploi	28/03/2017	39183 (Moyenne = 125 ; Médiane = 0)	5	3	- Application (1) - Visualisations (2)	Emploi et formation	- Moyenne : 6137 - Médiane : 104 - Min : 17 - Max : 18290

Type	Exemple examiné	Format	Emetteur	Usage	Cible	Structure et contenu	Canal	Pratiques intéressantes
Synthèse de cas d'usage	A quoi servent les données publiques ? Mise en perspective des meilleurs cas d'usages des données publiques ouvertes	Rapport PDF	Open Data France	Communiquer sur les usages de l'open data Convaincre de la pertinence de l'open data Valoriser les cas d'usage	Décideurs Grand public Producteurs de données	Structuration du rapport par thématiques 3 trames de présentation : - Pour un usage "générique" : 1/ Contexte-Besoin 2/ Liste de réutilisations listées en illustration et pas développées - Pour une solution spécifique : 1/ Contexte-Objectif 2/ Description de la solution 3/ Limites-besoins et perspectives en termes d'open data - Pour un usage d'un JDD particulier : 1/ Description du JDD 2/ Exemple concret d'utilisation 3/ Historique de l'ouverture du JDD 4/ Liste de réutilisations en illustration pas développées 5/ Limites et freins	Site web	Entretiens sur certaines réutilisations/cas d'usages : - Questions posées : Qu'est-ce que le projet ? Quelles données publiques sont utilisées et comment sont-elles utilisées ? Quel bilan et perspectives ? - Les entretiens concernant à la fois des producteurs et réutilisateurs de données Organisation des cas d'usages en dossiers thématiques
Synthèse de cas d'usage	Rapport "Pour une politique publique de la donnée, des algorithmes et des codes sources"	Rapport PDF	Mission Bothorel	Communiquer sur les usages de l'open data Convaincre de la pertinence de l'open data Faire l'état des lieux de l'open data Mobiliser les producteurs de données autour des limites en termes d'open data	Décideurs Equipes opérationnelles open data Producteurs de données	Cas d'usage viennent en illustration de chaque partie du rapport Structure pour un cas d'usage sur un JDD : 1/ Description du JDD 2/ Historique du JDD 3/ Utilisateurs du JDD 4/ Exemples d'usage du JDD 5/ Problèmes/freins et pistes d'améliorations pour le JDD	Circulation interne Réseaux sociaux Site web	Problématisation des cas d'usage Une partie "Enseignements à tirer"

Synthèse de cas d'usage	L'open data, ça sert à quoi ?	Article	Openda- tasoftware	Communi- quer sur les usages de l'open data Convaincre de la per- tinance de l'open data	Grand pu- blic Producteurs de données	Les cas d'usage, décrits de manière concise, viennent illustrer les impacts de l'open data	Blog	Organisation par type d'impact de l'open data et sous organisa- tion par thématique
Synthèse de cas d'usage	Open Data's Impact	Rapport PDF	GovLab	Communi- quer sur les usages de l'open data Convaincre de la per- tinance de l'open data Inspirer les réutilisateurs Partager de bonnes pratiques en termes d'open data	Décideurs Grand pu- blic Producteurs de données Réutilisa- teurs de données	Cas d'usage classés par type d'impact de l'open data : - Improving government - Empowering citizens - Creating opportunity - Solving public problems 1 pdf par cas d'usage Structure du cas d'usage : Une page «Summary» avec rappel du contexte, «Dimensions of impact» et «Key takeaways» I) Context and background II) Project description and inception III) Impact IV) Challenges V) Looking forward Parfois aussi une section «Lessons learned» OU I) The case in numbers II) Background III) The data IV) The path to impact V) Impact VI) Discussion VII) Calls to action	Site web	- Un système d'open peer review pour nour- rir les cas d'usage - Inclut des citations

Portail	Site d'Open-DataFrance	Page éditorialisée	OpenDataFrance	Communiquer sur les usages de l'open data Faire un inventaire des réutilisations Valoriser les réutilisations	Grand public Producteurs de données Réutilisateurs de données	Page «Réutilisations» qui constitue une galerie de réutilisations sélectionnées, classées par thème Page d'une réutilisation - Description de la réutilisation - Lien vers la réutilisation - Type de réutilisation - Réutilisateur - Profil du réutilisateur - Données utilisées - Thèmes - Mots clés	Site web	
Portail	Portail open data du Canada	Page éditorialisée	Canada	Inspirer les réutilisateurs Valoriser les réutilisateurs Valoriser les réutilisations	Grand public Réutilisateurs de données	Page «Galerie d'applications» - Distinction entre applications créées par l'administration vs les autres - Possibilité de noter les applications et de les commenter - Page de chaque application : Courte description, lien pour télécharger l'application, concepteur, jeux de données utilisées, sujet, type d'appareil, date de publication, mots clés Page "Histoire des utilisateurs de données ouvertes" - Histoires sélectionnées (au nombre de 17), liste non catégorisée - Possibilité de liker et de laisser un commentaire - Structure d'une histoire : "Créé par", "Site Web", "Sujet", "Résumé", "Impact", "Témoignages"(citation), "Jeux de données utilisés" et recommandation d'autres histoires	Portail	- Système de recommandations d'autres histoires - Possibilité de liker et de commenter

Portail	Portail open data de l'Italie	Article	Italie	Valoriser les cas d'usage	Grand public Réutilisateurs de données		Blog	Un tag "Use case" sur les articles de blog
Portail	Portail open data de New York	Page éditorialisée	New York City	Inspirer les réutilisateurs Valoriser les réutilisateurs Valoriser les réutilisations	Grand public Réutilisateurs de données	Page «projects» - Sélection de réutilisations soumises par les réutilisateurs - Catégories : «Interesting tech», data journalism», «civic advocate», «map», «research lab», ... - Tags Page de la réutilisation : - "How I started" - "Used data" - "How I built this" Courte biographie du réutilisateur	Portail	- Biographie du réutilisateur
Portail	Portail open data de Toronto	Page éditorialisée	Toronto	Valoriser les réutilisations	Grand public Réutilisateurs de données	Page réutilisation - Sélection de réutilisations, pas de page pour chaque réutilisation mais un lien direct vers la reuse	Portail	

Portail	Portail open data de Toronto	Tutoriel	Toronto	Former les réutilisateurs Inspirer les réutilisateurs	Grand public Réutilisateurs de données	Section "Data stories" sur le blog Tutoriels (exemple de structure : Define research questions, Prepare data for analysis, Visualise data for exploration, Analyse patterns and outliers, Engage with data owners)	Blog	Le format de tutoriel sur un cas d'usage
Portail	Portail open data des Pays Bas	Catalogue de réutilisations	Pays Bas	Faire un inventaire des réutilisations Valoriser les réutilisations	Grand public Réutilisateurs de données	Section «Impact of data» avec la sous catégorie «Applications with data» Page réutilisation : Description du projet + Détails sur le réutilisateur	Portail	
Portail	Portail open data des Pays Bas	Interview	Pays Bas	Inspirer les réutilisateurs Valoriser les réutilisateurs Valoriser les réutilisations	Grand public Réutilisateurs de données	Dans la partie «News», des «Impact Story» Questions des interviews : Can you briefly introduce yourself? Who are you and where do you work? What application have you worked on?Who did you do that for and why?What problem do you solve with that?What does this application mean for the user?What has changed for your users since they have this application?What data do you use for this?Every project also has challenges. What did you run into?If you look back, what has turned out differently than you thought?What is your personal connection to the subject?What's next? What's on the agenda in terms of data?What can Data.Overheid.nl do for you? What do you need?Is there anything else you would like to say about this topic?If people need more information after reading it, where can they turn?	Portail	Le format d'interview qui met en valeur le réutilisateur et met en récit la démarche —> La réutilisation est incarnée

Portail	Portail open data de la Colombie	Catalogue de réutilisations	Colombie	Faire un inventaire des réutilisations Valoriser les réutilisations	Grand public Réutilisateurs de données	Catalogue de réutilisations à parcourir : - Organisations - Thématiques - Type - Tags Page de la réutilisation - Que la réutilisation sans description	Portail	
Portail	Portail open data de l'Espagne	Catalogue de réutilisations	Espagne	Faire un inventaire des réutilisations Valoriser les réutilisations	Grand public Réutilisateurs de données	Une section «Impact» avec : - Reuse companies Catalogue des reuse companies par thématique, tags Page reuse companies : Thématique/Date de création/Page web/Résumé/Main products and services/Key to success/Growth/Future plans/Data used/Customers/Business model/Contact - Applications Catalogue par thématique, plateforme, tags Page reuse : date de publication, url, réutilisateur, résumé, images de l'app, tags, contenu associé, commentaires	Portail	Catalogue des reuse companies

Portail	Portail open data du Grand Lyon	Article	Grand Lyon	Communi- quer sur les usages de l'open data Inspirer les réutilisateurs Valoriser les réutilisations	Grand pu- blic Réutilisa- teurs de données	- Articles éditorialisation des réutilisations (listing) - Articles mise en récit d'une réutilisation	Blog	
Portail	London Data Store		Royaume Uni	Communi- quer sur les usages de l'open data Rendre l'information accessible	Grand pu- blic Réutilisa- teurs de données	Une série de dataviz sur la page d'accueil, par théma- tique	Portail	Des éditorialisations in- téressantes de données disponibles dès l'entrée sur le portail
Portail	London Data Store	Article	Royaume Uni	Communi- quer sur les usages de l'open data Rendre l'information accessible Valoriser les réutilisations	Grand pu- blic Réutilisa- teurs de données	- Articles avec des analyses de données qui délivre les insights principaux - Articles mise en récit d'une réutilisation	Blog	
Portail	Portail eu- ropéen des données	Cata- logue de réutilisa- tions	Union euro- péenne	Faire un inventaire des réutilisations Valoriser les réutilisations	Grand pu- blic Réutilisa- teurs de données	Page "Réutilisations" - Catalogue de réutilisa- tions - Fiche réutilisation (pdf) : URL, Quick Facts, Descrip- tion, Benefits, How open data is used	Portail	

Solution de gestion et partage de données	data.world	Page éditorialisée Rapport PDF	data.world	Convaincre de la pertinence des services proposés Promouvoir ses services avec une vitrine Valoriser les cas d'usage	Clients potentiels	Section «Case Study» > Version site - With data.world collaboration, they: - The challenge - The solution - Results - What's next ? Citation Recommandations d'autres ressources utiles > Version pdf - Description du client - Challenge - Solution - Results - Challenges	Site web	
Solution de gestion et partage de données	Dawex	Rapport PDF	Dawex	Convaincre de la pertinence des services proposés Promouvoir ses services avec une vitrine Valoriser les cas d'usage	Clients potentiels	Structure : - Meet Agdatahub - The solution - The result - The opportunity	Site web	Il faut se référencer pour télécharger une étude de cas (format pdf) —> identifier des acteurs intéressés

Solution de gestion et partage de données	Kaggle Winner’s blog	Tutoriel	Kaggle	Former les réutilisateurs Inspirer les réutilisateurs Valoriser les réutilisateurs Valoriser les réutilisations	Réutilisateurs de données		Blog	Le format de tutoriel
Solution de gestion et partage de données	Kaggle Winner’s blog	Interview	Kaggle	Inspirer les réutilisateurs Valoriser les réutilisateurs Valoriser les réutilisations	Réutilisateurs de données	Structure et questions : 1/Let’s learn about David - David, what can you tell us about your background? - Do you have any prior experience or domain knowledge that helped you succeed in this competition? - How did you get started competing on Kaggle? What made you decide to enter this competition? 2/Let’s get technical - Tell us about the overall architecture or approach to the problem. - Did any past research or previous competitions inform your approach? - What preprocessing and feature engineering did you do? - What supervised learning methods did you use? - What was your most important insight into the data? - Were you surprised by any of your findings? - Which tools did you use? - How did you spend your time on this competition? - What does your hardware setup look like? - What was the run time for both training and prediction of your winning solution? 3/Words of Wisdom - What have you taken away from this competition? - Looking back, what would you do differently now? - Do you have any advice for those just getting started in data science? 4/Just for fun	Blog	

Solution de gestion et partage de données	Site d'Open-datasoft	Rapport PDF	Open-datasoft	Communiquer sur les usages de l'open data Convaincre de la pertinence de l'open data Promouvoir ses services avec une vitrine Valoriser les cas d'usage	Clients potentiels	"Histoires client" Courte description —> Se référencer pour télécharger les éléments (pdf) Structure : - Le challenge - La solution - Résultats clés - L'impact	Site web	
Solution de gestion et partage de données	Site d'Open-datasoft	Vidéo	Open-datasoft	Communiquer sur les usages de l'open data Convaincre de la pertinence de l'open data Valoriser les cas d'usage	Clients potentiels		Site web	
Solution de gestion et partage de données	Open-datasoft Academy	Catalogue de réutilisations	Open-datasoft	Faire un inventaire des réutilisations Inspirer les réutilisateurs Valoriser les réutilisations	Grand public Producteurs de données Réutilisateurs de données	Catalogue avec directement : - Organisation - Courte description - Objectif - Lien vers l'initiative	Site web	

Autres	Géoservices	Page éditorialisée	IGN	Communiquer sur les usages de l'open data Inspirer les réutilisateurs	Grand public Réutilisateurs de données	Page «Usages» avec toutes les thématiques Pour chaque thématique, une page éditorialisée avec 3/4 cas d'usage présentés, 1 témoignage et une suggestion des données et services à disposition	Portail	Le format de la page
Autres	Hackers publics	Podcast	Camille Boulouis	Communiquer sur les projets Inspirer de nouveaux projets Valoriser les porteurs de projets Valoriser les projets	Agents publics	Avec Hackers publics, partez à la rencontre celles et ceux qui transforment la culture numérique de l'administration. Prenez une grande bouffée d'inspiration et découvrez l'avant-garde de l'innovation publique. Si vous êtes dans l'administration, vous verrez que vous n'êtes pas seul à essayer de changer les choses. Si vous n'en faites pas (encore) partie, vous découvrirez qu'il est possible d'innover tout en ayant un travail qui a du sens ! Et qui sait, ça vous donnera peut-être envie de nous rejoindre pour nous donner un coup de main ?		
Autres	Datagora	Infographie Vidéo	Datagora	Rendre l'information accessible, Valoriser les études	Grand public	Infographie ou dataviz accompagnée d'une courte phrase qui énonce l'insight principal Vidéo : chiffres clés par thématique/sujet, contextualisés	Site web	Format synthétique sur un sujet

Autres	Datagora	Infographie Vidéo	Datagora	Rendre l'information accessible Valoriser les études	Grand public	Vidéo : Décryptage avec une personne qui présente Infographie : Thread par sujet - infographie = Phrase avec l'insight principal + une dataviz	Réseaux sociaux	
Autres	api.gouv.fr : Créer un moteur de recherche des associations françaises	Guide	api.gouv.fr	Former les réutilisateurs, Inspirer les réutilisateurs	Réutilisateurs de données	1/ Qu'est ce que l'API Répertoire National des Associations (RNA) ? 2/ Cas pratique : rechercher une association par son nom 3/Comment ça marche ? 4/ Aller plus loin		