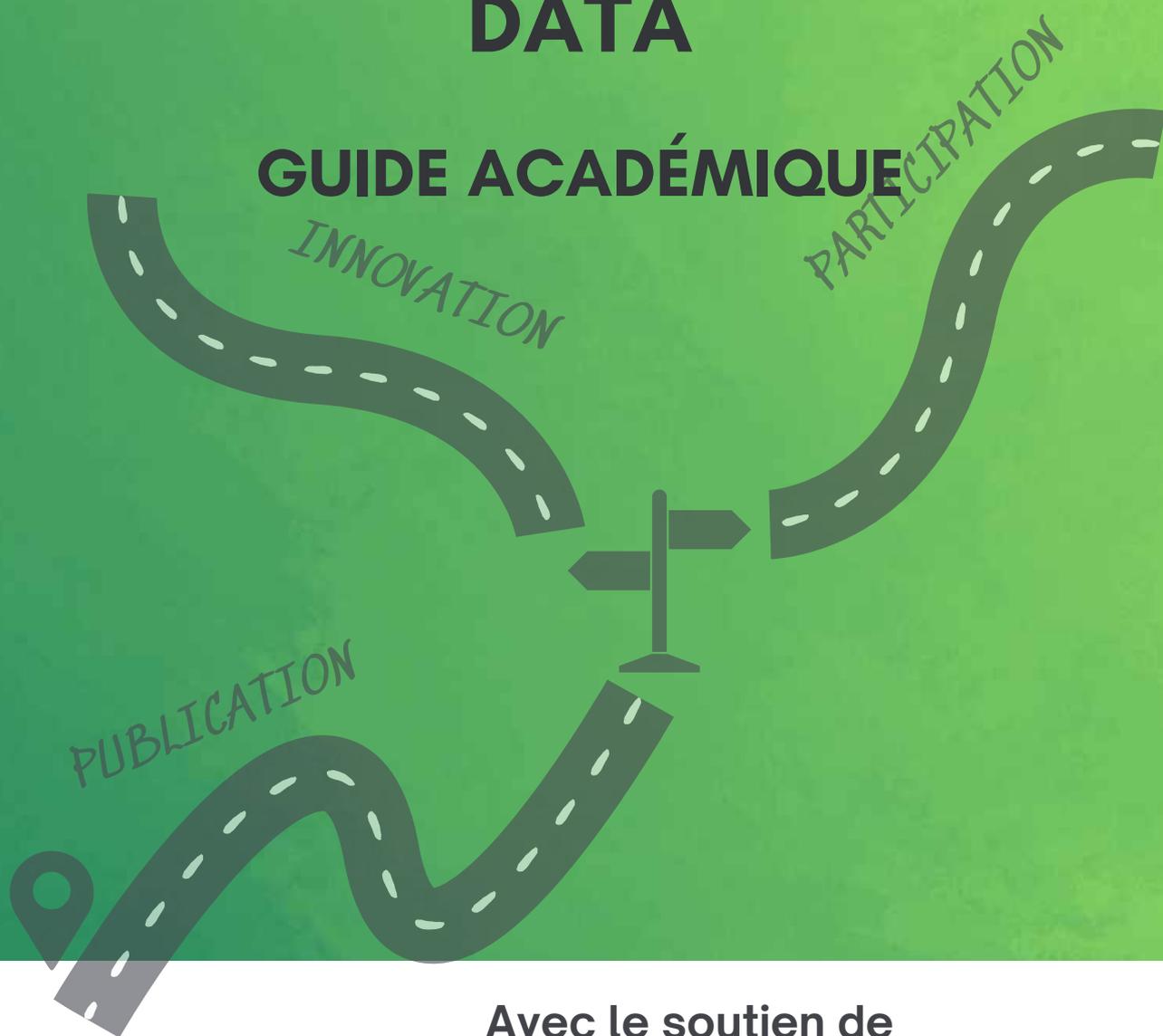


ENGAGER LES RÉUTILISATEURS DE LA PUBLICATION À L'EXPLOITATION DE L'OPEN DATA

GUIDE ACADÉMIQUE



Avec le soutien de



Table des matières

- 1** Résumé
- 2** La recherche en open data au NADI
- 3** Introduction
- 4** Route de la publication
- 5** Route de l'innovation
- 6** Route de la participation
- 7** Conclusion
- 8** Références scientifiques

Résumé

Ce guide a été rédigé par le Namur Digital Institute dans le cadre de sa mission de Référent académique Smart Region. Il présente le processus de publication d'open data, ainsi que sa réutilisation au travers de l'innovation et de la participation citoyenne. Il en passe en revue les grandes étapes et défis et propose des recommandations basées sur la littérature scientifique, et en particulier les recherches appliquées menées par le Namur Digital Institute en Wallonie. Il se veut être une ressource complémentaire aux documents et services développés par les autres référents, en particulier FuturoCité et le Smart City Institute.

Le guide est tant à destination des entités publiques souhaitant se lancer dans la publication d'open data que celles déjà engagées dans le processus mais souhaitant s'inspirer du regard scientifique et d'exemples en la matière. Il passe d'abord en revue le processus de publication d'open data et propose 15 recommandations pratiques basées sur la littérature scientifique. Le guide présente ensuite en détails comment l'open data peut être réutilisé pour favoriser l'innovation et la participation citoyenne. Loin de traiter la publication et la réutilisation d'open data comme des processus indépendants, ce guide prend le positionnement inverse et explique comment ces processus sont intrinsèquement liés, et ce qui peut être fait lors de la publication pour favoriser la réutilisation d'open data.

Au travers de ce guide, nous voulons donc essentiellement transmettre que favoriser la réutilisation de l'open data se réfléchit dès sa publication. Notre message clé est qu'un processus de publication d'open data centré sur les réutilisateurs de l'open data, c'est-à-dire un processus de publication dans lequel il est nécessaire de comprendre et s'engager avec eux, est essentiel à la réutilisation.

La recherche en open data au NADI

Ces dernières années, l'évolution des technologies a permis à de nombreuses organisations de fournir des informations et des services d'une manière innovante. En ce qui concerne les gouvernements, ces nouveaux développements ont conduit à l'émergence du concept d'"e-gouvernement", qui fait référence à l'utilisation des technologies par les entités gouvernementales, afin de fournir leurs informations et leurs services de manière plus optimale aux entreprises, citoyens, ainsi qu'aux autres organisations gouvernementales. Le NADI - **Namur Digital Institute** - promeut la recherche scientifique dans ce domaines sous différents angles : simplification administrative, exigences en matière de sécurité, développement des services en ligne, participation citoyenne, interaction humain-machine, mais aussi l'open data.

Le NADI a une vision de la recherche très axée sur l'interdisciplinarité (technologie, impacts sociétaux et managériaux, cadre légal, etc.). A l'image de cette vision, la recherche sur l'open data au NADI se veut elle aussi interdisciplinaire. Fort de plusieurs années d'expérience de recherche dans ce domaine, le NADI encourage les collaborations entre chercheurs possédant des expertises différentes, que ce soit en son sein ou au-delà, et avec les acteurs de terrain, que ce soit du côté des publieurs ou des réutilisateurs d'open data. Les recherches du NADI sur l'open data ont été publiées dans de nombreuses revues et conférences scientifiques de rang international, permettant au NADI d'être aujourd'hui un expert reconnu internationalement dans le domaine.

La recherche sur l'open data au NADI porte sur nombre d'aspects différents, depuis l'analyse des cadres légaux pertinents (Directive Open Data, RGPD, etc.), jusqu'à l'étude des bonnes pratiques favorisant la réutilisation de l'open data, en passant par des considérations plus techniques telles que les formats de données et la qualité des données. Les chercheurs du NADI réalisent tant des études de la littérature scientifique existante que des études de terrain, impliquant des citoyens, entités publiques, développeurs de services, et d'autres.

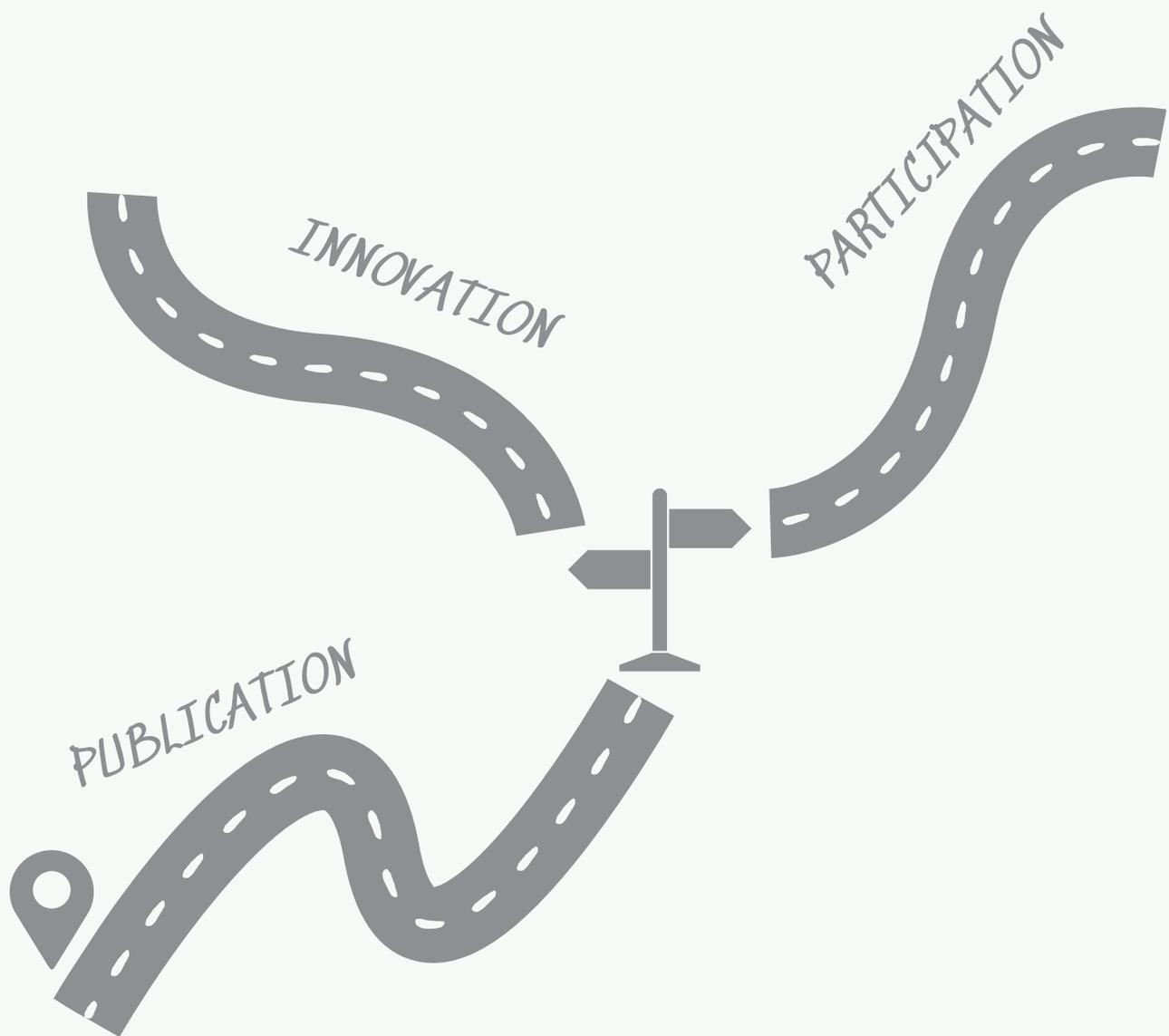
Introduction

Les données ouvertes, ou l'open data, sont des données interopérables publiées sur Internet afin qu'elles puissent être librement consultées et réutilisées par tous ceux qui le souhaitent (Attard et al., 2015). Nous nous intéresserons dans ce guide à un type d'open data particulier, à savoir celui publié par les entités publiques et visé par la Directive Open Data de l'Union Européenne (Directive (EU) 2019/1024). Ces données peuvent concerner des aspects divers et variés du territoire de l'entité publique et porter sur, par exemple, les montants alloués à chaque poste du budget, les places de parking disponibles, la démographie par quartier, ou encore le petit patrimoine.

La publication en ligne d'open data se fait habituellement en préparant des jeux de données et en les mettant à disposition via une plateforme en ligne accessible à tous. La publication d'open data est un processus comportant plusieurs étapes, dont l'élaboration d'une stratégie de publication et la documentation des jeux de données, entre autres (Kucera et al., 2015 ; Crusoe et al., 2020a). Dans ce guide, nous passerons en revue les grandes étapes de cette **Route de la publication**. Nous mettrons quelques recommandations pratiques basées sur la littérature scientifique et sur les recherches du NADI menées dans le contexte wallon, ainsi que des illustrations de certaines de ces recommandations. Enfin, nous montrerons comment suivre ces recommandations contribue aux indicateurs de la Smart Région définis par le Smart City Institute.

La littérature scientifique et les textes légaux mettent en avant plusieurs bénéfices qui peuvent être réalisés si les entités publiques publient de l'open data de qualité suffisante. Ces données peuvent être réutilisées par le secteur privé pour développer des services innovants (Janssen et al., 2012 ; Attard et al., 2015 ; Directive (EU) 2019/1024) ou mieux cibler son marché, par exemple grâce à l'open data sur la démographie de la population (Magalhaes & Roseira, 2020). La société civile peut également bénéficier de l'open data au travers de l'utilisation de services développés grâce à ces données (Chokki et al., 2022a). De plus, la publication de l'open data peut améliorer la transparence de l'action publique en donnant aux citoyens de l'information sur l'état du territoire et les activités des entités publiques qui le gèrent (Ruijter et al., 2017).

Enfin, l'open data peut être exploité de plusieurs manières différentes dans des formes de participation citoyenne plus actives, où la donnée peut servir de support pour permettre aux citoyens de s'exprimer dans les meilleures conditions (Bvuma & Joseph, 2019 ; Kempeneer & Wolswinkel, 2023 ; Ruijer et al., 2024). Nous passerons en revue les directions que peuvent ainsi prendre la **Route de l'innovation** et la **Route de la participation**, en présentant de nombreux exemples de réutilisation et en proposant quelques recommandations pour aider les entités à favoriser la réutilisation de leurs données.



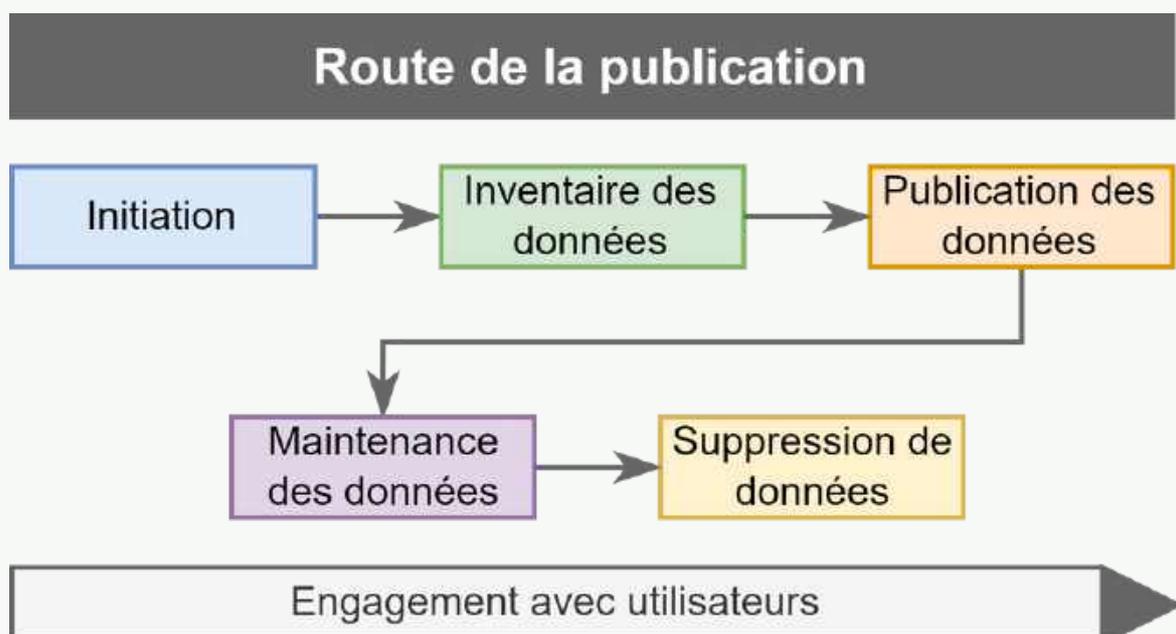
Route de la publication

Aperçu du chapitre

Ce chapitre présente le processus de publication d'open data par une entité publique tel que décrit dans la littérature scientifique. Plusieurs recommandations pratiques sont formulées et illustrées aux différentes étapes du processus. Le chapitre montre enfin le lien entre les recommandations et les indicateurs de la Smart Région définis par le Smart City Institute.

Processus de publication d'open data et recommandations

La publication d'open data par une entité publique est un processus comportant plusieurs étapes. Dans la littérature scientifique (Crusoe et al., 2020a), il est décrit en six étapes, allant de l'initiation du projet de publication d'open data jusqu'à l'engagement avec les utilisateurs. Nous le modifions légèrement car nous prenons le parti d'un engagement avec les utilisateurs transversal au processus dans son ensemble. Afin d'assurer que ce processus est mis en place de la meilleure manière possible, il est recommandé au lecteur de ce guide de consulter le chapitre 2 du Guide Pratique sur la gouvernance de la donnée du [Smart City Institute \(2019, p. 34\)](#). Il constitue une ressource complémentaire guidant la mise en place d'une stratégie de gouvernance de la donnée.



Processus de publication d'open data adapté de (Crusoe et al., 2020a).

L'étape d'**initiation** donne l'impulsion du projet de publication d'open data, ou en d'autres termes la raison pour laquelle une entité publique commence à publier de l'open data. La raison peut varier (Nikiforova et al., 2024), une entité peut par exemple initier une publication d'open data en prenant de l'inspiration chez une autre entité. L'entité publique peut aussi publier de l'open data pour respecter l'obligation légale. L'entité peut aussi se projeter dans le type de réutilisation qu'elle souhaite susciter en priorité. Souhaite-t-elle par exemple favoriser en priorité la transparence de son action ou la création de services innovants ? Il est essentiel de bien identifier et s'accorder sur les objectifs de la publication d'open data car ils vont affecter les bonnes pratiques à suivre lors des autres étapes du processus de publication.



Recommandation 1

Au sein de l'entité, formaliser et s'accorder sur les objectifs de la publication d'open data.

Il est également important que les personnes impliquées dans le projet de publication d'open data en comprennent bien l'intérêt, afin d'éviter des blocages dans le processus. Dès lors, il faut organiser des discussions en interne dans l'entité, encourageant chacun à exprimer son point de vue. Il est aussi essentiel de sensibiliser les membres de l'entité à l'intérêt de l'open data, par exemple en mettant en avant des retours d'expérience d'autres entités.



Recommandation 2

Dès le début du projet de publication, chercher à sensibiliser les membres de l'entité à l'intérêt de l'open data tout en écoutant les points de vue de chacun.

Déjà lors de l'étape d'initiation, il peut être opportun de prendre contact avec d'autres entités envisageant également de publier de l'open data. Cela permet un échange de bonnes pratiques lors des activités de formalisation des objectifs et de sensibilisation. Cela permet également d'identifier les entités ayant un objectif de publication similaire en vue d'engager des collaborations lors de la réutilisation des données. Par exemple, une coordination de plusieurs entités pour fournir un service innovant à l'échelle de plusieurs territoires, ou encore une participation citoyenne coordonnée.

Lors de l'étape d'**inventaire des données**, l'entité publique prend note des données qu'elle possède et qu'elle doit ou peut rendre publiques selon les cadres légaux. Lors de cet inventaire, il est important de ne pas se limiter aux données déjà présentes dans les outils numériques. En effet, le concept de "donnée" est très large. Au-delà des données numériques dont une liste d'exemples est fournie dans le Guide Pratique sur la gouvernance de la donnée du [Smart City Institute \(2019, p. 23\)](#), il est possible que certaines données existent sous format papier au moment de l'inventaire. Il convient alors d'effectuer cet inventaire en impliquant suffisamment de personnes de l'entité afin de s'assurer d'avoir une vue complète des données existantes. En effet, ce qui est considéré comme de la donnée pour une personne (par exemple, des documents dans un classeur) peut ne pas l'être pour une autre personne.



Recommandation 3

Impliquer le plus grand nombre de membres de l'entité lors de l'inventaire des données.

Au-delà du défi d'identifier les données existantes en interne, il peut arriver que des données que l'entité souhaiterait publier en open data n'existent pas. Dans ce cas, elle doit produire la donnée. Elle dispose de deux possibilités pour le faire. Premièrement, si les données sont géolocalisées, elle peut lancer un appel auprès de la population et demander aux citoyens qui sont sur le territoire de faire remonter les données auprès d'elle, dans la logique du citoyen comme capteur ([Goodchild, 2007](#)). Nous reviendrons plus en détail sur ce cas de figure dans le chapitre sur la Route de la participation citoyenne. Deuxièmement, l'entité peut aller elle-même collecter les données. Nous recommandons dans ce cas, afin de diminuer les coûts de la démarche, de mutualiser tant que possible l'infrastructure existante.



SERVICE

Camions-sentinelles

La flotte de camions du Bureau Economique de la Province (BEP) parcourt une à deux fois par semaine l'ensemble des routes de la Province de Namur. Depuis longtemps, le BEP cherchait à **valoriser ces kilomètres parcourus chaque jour par sa centaine de camions**. A l'heure où la donnée est un bien de plus en plus précieux et que la connaissance du territoire est indispensable pour continuer à offrir les services les plus pertinents qui soient aux Communes, l'idée s'est imposée : les camions poubelles deviendront camions-sentinelles.

Désormais aux missions de collecte, traitement et valorisation des déchets seront ajoutées la collecte, le traitement et la valorisation de données. Concrètement, le projet consiste à équiper 40 camions poubelles d'un hub de capteurs, boîtier sur lequel n'importe quel type de capteurs peut venir se connecter afin de collecter des données sur l'ensemble du parcours de chaque camions. Ce boîtier est capable de prétraiter les données collectées avant transfert vers une plateforme développée spécifiquement pour gérer les données enregistrées.



*Texte extrait du [site du BEP](#)
Capture d'écran : [Infopole Wallonie](#)*

Ensuite, l'entité priorise les jeux de données qu'elle va mettre à disposition en premier. En effet, la préparation d'un jeu de données pour publication peut s'avérer chronophage dans certains cas. Les données peuvent devoir être entièrement numérisées, ou converties dans un format ouvert car elles sont issues d'un outil propriétaire. Elles peuvent aussi comporter des problèmes de qualité ou d'incomplétude qui doivent être résolus avant la publication. Enfin, les cadres légaux peuvent poser des contraintes à ce stade. Par exemple, les données à haute valeur (High Value Datasets) telles que définies dans la Directive Open Data (c'est-à-dire les données géospatiales, les données relatives à l'observation de la Terre et à l'environnement, les données météorologiques, les statistiques, les données relatives aux entreprises, et les données relatives à la mobilité) doivent être disponibles en format lisible par ordinateur (tel que CSV ou JSON) et être disponibles via une API (Application Programming Interface). Un autre exemple est l'anonymisation des données à caractère personnel avant leur publication, tel qu'exigé par le RGPD. La littérature scientifique recommande aux entités ne disposant pas du suffisamment d'expertise dans les cadres légaux liés à l'open data de consulter des experts.

Une explication détaillée des différents textes légaux pertinents pour l'open data ne figure pas dans ce guide. Cependant, **dans sa mission de référent académique Smart Région, le NADI met à disposition une expertise légale liée à l'open data.**



Recommandation 4

Si l'expertise légale manque en interne, solliciter l'avis d'experts en cas de doute avant la publication des données.

Ces différents éléments font qu'il peut être plus ou moins facile de rendre un jeu de données prêt à la publication, et qu'il est impossible de tout publier en même temps. L'entité devra donc décider de quelques jeux de données à ouvrir en premier lieu. Certains critères de décision bien connus incluent l'expérience personnelle des personnes en charge de la publication des données, la disponibilité des données en interne, les requêtes par d'autres membres de l'entité, ou encore la comparaison avec ce qui est publié sur les portails d'autres villes (Yang & Wu, 2021). Une autre possibilité est de commencer par des jeux de données plus faciles pour se faire la main progressivement ou de choisir quelques jeux de données autour desquels l'entité souhaite réaliser un projet nécessitant une ouverture vers l'extérieur.



Recommandation 5

Commencer par publier des jeux de données dont la préparation est moins difficile, afin de se faire la main progressivement.

Un autre critère de priorisation peut se trouver dans le public que l'entité souhaite le plus cibler. La littérature scientifique met en avant que le grand public et ceux qui vont utiliser l'open data pour développer un service innovant ont des attentes différentes quant aux données publiées (Simonofski et al., 2022). Les développeurs de services innovants souhaitent des données exploitables et lisibles par ordinateur, sans préférence pour un thème ou l'autre. Si c'est le public principal visé par une entité, les thématiques mises en avant dans les High Value Datasets sont considérées par la Directive Open Data comme particulièrement intéressantes dans ce cas de figure. D'un autre côté, le grand public a une préférence plus prononcée pour les données de transparence telles que les budgets de l'entité publique.



Recommandation 6

Si un public cible principal pour l'open data est clairement identifié, en tenir compte pour la priorisation des jeux de données. Pour le grand public, préférer les données de transparence, et pour les développeurs de services les High Value Datasets.

Bien qu'ils aient le mérite d'être applicables à différentes entités, et qu'ils permettent de prendre des décisions de priorisation raisonnées, il reste possible qu'il y ait un problème d'adéquation entre l'offre et la demande d'open data. En d'autres termes, les jeux de données disponibles ne sont pas réutilisés car ils ne correspondent pas à ce que la population souhaite voir. Dès lors, de nombreux scientifiques recommandent aux entités publiques de s'intéresser à l'utilisateur cible des jeux de données en cherchant à déterminer de quelles données ils souhaitent disposer (Janssen et al., 2012 ; Ubaldi, 2013 ; Dawes, 2016 ; Ojo, 2016 ; Hogan, 2017). Une manière d'étudier la question est de récolter leurs besoins en informations (Wilson, 1981). Des approches comme les ateliers (Crusoe et al., 2020b) peuvent être organisées avec des groupes de citoyens pour mieux comprendre de quelles informations ils souhaiteraient disposer. Toutefois, d'autres approches existent, dont les avantages et inconvénients sont résumés ci-dessous. Dans l'ensemble, ces approches demandent des ressources et du temps mais ont l'avantage d'aider à comprendre les besoins en information spécifiques des citoyens du territoire.

Recommandation 7



Si les jeux de données sont moins utilisés qu'espéré, et qu'aucun autre problème évident n'empêche la réutilisation des données, vérifier qu'il n'y a pas d'inadéquation entre l'offre et la demande d'open data.



Recherche de logement

PROTOTYPE

Dans le cadre du cours de projet en Data Analytics, des étudiants de l'Université de Namur ont développé **LogezVous**, un service visant à aider les citoyens à choisir une zone dans Namur où chercher un logement sur base de certains critères (transports, santé, éducation, culture, sport). Le service leur renseigne, en fonction des préférences qu'ils ont encodées, une zone d'habitation qui correspond le mieux à leurs critères.

Cet exemple montre une chose importante. Pour que ce service soit le meilleur possible, il faut une connaissance précise de ce que les citoyens recherchent comment information quand ils choisissent un endroit où habiter. Dès lors, une bonne **connaissance de leurs besoins en information** permet de publier les jeux de données nécessaires à la création du service qui sera le plus utile possible.



Résumé des avantages et inconvénients des approches d'étude des besoins en information.

Approche	Avantages	Inconvénients
Questionnaire	<ul style="list-style-type: none"> Grand nombre de répondants Possibilité de faire des analyses statistiques 	<ul style="list-style-type: none"> Ne capture pas toute la complexité des besoins en information
Entretiens avec des citoyens individuels	<ul style="list-style-type: none"> Possibilité de rebondir sur ce que le citoyen dit Plus facile à organiser au niveau des disponibilités 	<ul style="list-style-type: none"> Peu de répondants Les citoyens doivent être pris un à un Les questions sont définies à l'avance
Focus group	<ul style="list-style-type: none"> Rassemble plusieurs citoyens et amène à un consensus 	<ul style="list-style-type: none"> Les questions sont définies à l'avance Complexe à organiser (préparation d'outils, rassembler les citoyens)
Atelier participatif	<ul style="list-style-type: none"> Rassemble plusieurs citoyens et amène à un consensus Permet d'élaborer des solutions Ce sont les citoyens qui définissent les questions abordées 	<ul style="list-style-type: none"> Complexe à organiser (préparation d'outils, rassembler les citoyens) Nécessite une expertise en animation/facilitation

Lors de l'étape de **publication des données**, l'entité prépare les jeux de données afin qu'ils soient prêts à être rendus publics. Un aspect essentiel de cette préparation est de veiller à respecter, autant que possible, des standards existants dans les données. Même si cela peut demander du travail supplémentaire, il y a des avantages non négligeables.

Premièrement, cela permet une meilleure cohérence des jeux de données publiés par la même entité. La littérature scientifique a montré à maintes reprises que c'est lorsqu'elles sont combinées que les données atteignent leur plus grand potentiel. Par exemple, la Ville de Namur possède des données sur la population par quartier et le potentiel photovoltaïque par quartier. Si le quartier est encodé de manière différente dans les deux jeux de données (par exemple, une fois avec le nom et une fois avec un identifiant), il devient compliqué pour un réutilisateur de croiser les données par quartier (Crusoe et al., 2019). Deuxièmement, cela permet une meilleure cohérence et des croisements facilités avec d'autres jeux de données au-delà de l'entité publique. En effet, même s'ils concernent un territoire restreint, des jeux de données peuvent être utiles au-delà des limites du territoire. Par exemple, elles peuvent être utilisées dans un service qui aurait pour vocation d'intégrer les données de plusieurs territoires pour donner une vue d'ensemble et faire des comparaisons, ou encore combiner les données de plusieurs entités afin d'en avoir suffisamment que pour développer des modèles prédictifs robustes. Dès lors, se mettre d'accord entre entités sur des standards communs à adopter est essentiel pour favoriser la réutilisation de données à plus grande échelle.



PROTOTYPE

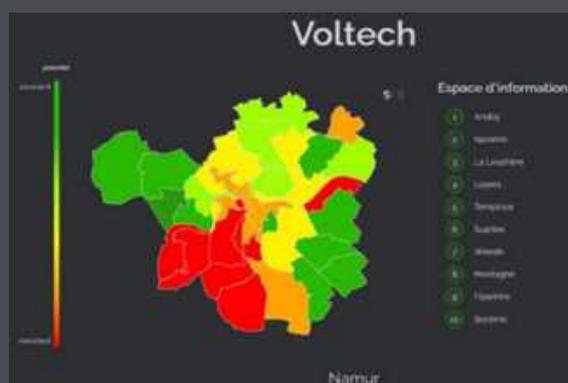


RETOUR D'XP

Potentiel photovoltaïque

Dans le cadre du cours de projet en Data Analytics, des étudiants de l'Université de Namur ont développé **Voltech**, un service visant à aider les citoyens dans leur décision d'installer des panneaux solaires. Le service est basé sur un algorithme de recommandation qui permet de visualiser le potentiel photovoltaïque d'un quartier ou d'un bâtiment de Namur. Le service utilise des jeux de données sur le potentiel photovoltaïque publiés en open data par la Ville de Namur. Ils ont cependant fait appel aux données de production des panneaux solaires disponibles en open data sur le portail de Londres afin d'améliorer la précision de leurs prédictions. Un autre groupe d'étudiants est également parti des données de la Ville de Namur sur le potentiel photovoltaïque et a utilisé les données météo pour affiner leurs prédictions.

Ces exemples illustrent que combiner les jeux de données de plusieurs entités publiques augmente incommensurablement le potentiel de réutilisation des données. Il est donc essentiel de faciliter la tâche des développeurs de services en veillant à utiliser des standards cohérents autant que possible quand on publie des jeux de données.



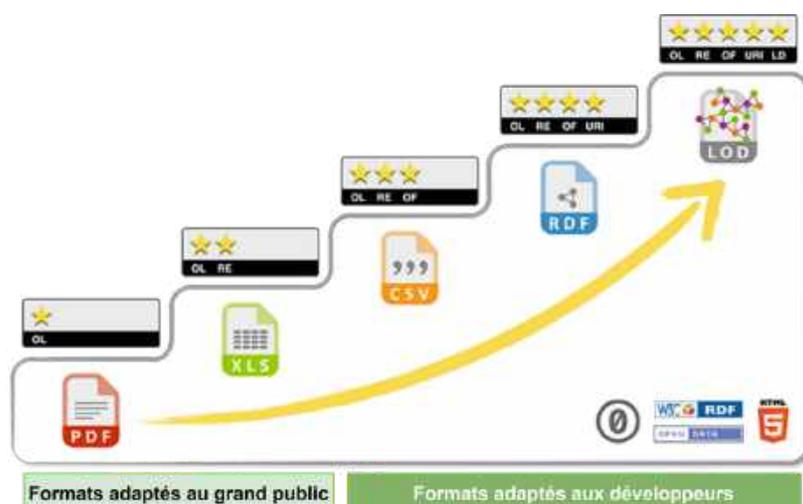
Pour reprendre l'exemple de **LogezVous**, il a été développé à partir des données de Namur, et donc ne permet aux citoyens de trouver une zone d'habitation que sur ce territoire. Développer ce service à partir des données d'une autre ville nécessitera un travail conséquent si les données sont représentées autrement (avec des noms de colonne différents, ou des attributs de données différents), mais ne nécessitera qu'un travail à la marge si les données sont standardisées de la même manière. Cela illustre qu'une collaboration entre entités dans la standardisation des open data favorise grandement la mutualisation des services développés pour une ville.

Recommandation 8



Travailler de concert avec d'autres entités afin que des données concernant les mêmes entités soient publiées dans un standard commun, et afin de favoriser l'échange de bonnes pratiques et standards en général.

Au-delà des standards, l'entité doit réfléchir à la forme et au format dans lequel rendre ses jeux de données disponibles. Le modèle de Tim Berners-Lee à 5 étoiles décrivant la maturité des formats d'open data encourage les formats lisibles par ordinateur. Cependant, ils ne sont pas adaptés pour le grand public. Bien que ces formats soient nécessaires pour que des services puissent être développés à partir de ces données, le grand public n'a pas les compétences techniques pour les traiter et préférera des formats lisibles par un humain (Simonofski et al., 2022), même s'ils atteignent un score de maturité moins élevé sur le modèle.



Modèle de maturité de l'open data de Tim Berners-Lee, avec indication des formats plus adaptés au grand public (PDF, formats lisibles par des applications propriétaires grand public telles que MS Excel) et ceux davantage adaptés aux développeurs (CSV, JSON, RDF, Linked Open Data).

De plus, le grand public ne se contentera pas des données dans ce format et souhaitera également disposer de descriptions vulgarisées du contenu des données ainsi que des visualisations simples d'accès. Par contre, ils ne seront pas intéressés par les descriptions techniques, pourtant nécessaires aux développeurs qui réutiliseront les données pour créer des services. On observe deux publics avec des attentes bien différentes quant à la manière dont les données doivent être rendues disponibles. Les entités souhaitant cibler les deux publics doivent prendre cela en compte quand elles mettent à disposition des jeux de données et proposer quelque chose de satisfaisant pour chaque public, peut-être dans deux interfaces distinctes.



Recommandation 9

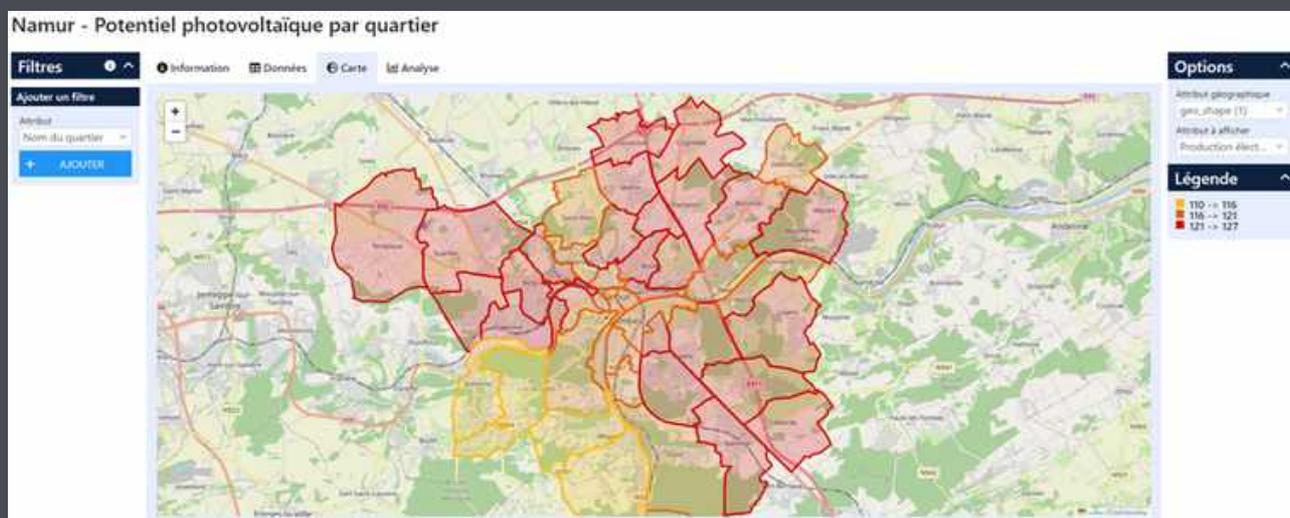
Tenir compte du public cible et de ses compétences techniques lors de la mise à disposition des données. Envisager une interface spécifique pour les développeurs et une autre pour le grand public.



PROTOTYPE

Visualisations interactives

Le grand public souhaitant explorer des données directement sur la plateforme où les données sont publiées n'a pas les compétences techniques pour traiter des données brutes et souhaitera des visualisations interactives lui permettant de prendre en main les données. Ainsi, pour être adaptée au grand public, l'offre d'open data doit inclure des visualisations des données telles que l'exemple ci-dessous, créé avec les données de potentiel photovoltaïque à Namur. Sur cette carte, les utilisateurs peuvent avoir une vue d'ensemble du potentiel photovoltaïque de chaque entité du territoire, et peut cliquer sur celle de son choix pour obtenir des informations plus détaillées.



Toutefois, ces deux publics présentent des points communs. Ils souhaitent tous les deux disposer de métadonnées décrivant chaque jeu de données et d'indicateurs de qualité des données tels que le nombre de données manquantes et la précision des données (Chokki et al., 2023). Ce genre d'informations documentant les données est important pour éviter que les données soient mal interprétées et mal utilisées par les réutilisateurs, ce qui pourrait amener à des conclusions erronées (Nikiforova et al., 2024). Les métadonnées peuvent par exemple décrire le contenu des données en général ou de colonnes plus spécifiquement, ou encore la date de mise à jour, la périodicité, et la source des données. Cette dernière est essentielle à renseigner si une entité publie des données issues d'une autre entité. Les réutilisateurs doivent en effet pouvoir retracer l'origine des jeux de données et évaluer leur fiabilité. Des mots-clés peuvent aussi être renseignés pour aider les réutilisateurs à trouver les jeux de données qui les intéressent parmi ceux publiés.



Recommandation 10

Documenter les jeux de données avec des descriptions et autres métadonnées pour éviter que les données soient mal interprétées par les réutilisateurs.



SERVICE

Cas de COVID-19

Des données sur les cas de COVID-19 à l'échelle nationale et locale sont disponibles sur le portail open data de la Ville de Namur. Afin d'aider les utilisateurs à interpréter correctement ces données, un onglet explique que les données de cas de COVID-19 au niveau national sont en cours de consolidation.



Capture d'écran : Portail open data de la Ville de Namur

A ce stade, l'entité publique dispose de jeux de données prêts à être publiés et sait comment les présenter aux utilisateurs. Elle a maintenant besoin d'une infrastructure lui permettant de mettre ces jeux de données en ligne. Plusieurs solutions sont possibles, telles que passer par (1) le site Internet de l'entité, (2) un portail open data public tel que ODWB, (3) un portail open data open source propre à l'entité, ou (4) un portail open data clé en main propre à l'entité. Le Guide Pratique sur la gouvernance de la donnée du [Smart City Institute \(2019, p. 71\)](#) fournit un excellent comparatif de ces quatre possibilités permettant de guider les entités.

Lors de cette étape de la publication, travailler de concert avec d'autres entités peut aussi s'avérer précieux. Par exemple, travailler en centrale d'achats ou collaborer dans l'écriture de cahiers de charge permet d'échanger de bonnes pratiques qui seront utiles dans d'autres situations, mais aussi d'économiser des ressources en les mutualisant.

Enfin, il est conseillé d'accompagner la publication de jeux de données avec des actions de sensibilisation. En effet, la littérature scientifique a depuis longtemps montré qu'il ne suffit pas de publier de l'open data pour garantir sa réutilisation (Ubaldi et al., 2013). Il faut aussi que les utilisateurs cibles soient au courant de l'existence des données et de leur valeur. Dès lors, une communication efficace autour de la publication de jeux de données et la mise en avant de réutilisations quand elles existent (même ultérieurement à la publication ou venant de l'intérieur de l'entité) permettent de sensibiliser les potentiels réutilisateurs à l'intérêt par l'open data. La nature de ces réutilisations peut changer en fonction de si c'est le grand public, les développeurs de services, ou les deux qui sont ciblés, en fonction de leurs différences. Par exemple, le grand public ayant un attrait plus prononcé pour les données de transparence, il sera plus judicieux de mettre en avant une réutilisation de type "portail de transparence" croisant plusieurs jeux de données sur le fonctionnement et les budgets de l'entité. Des thématiques pertinentes peuvent aussi être piochées en restant attentif aux sujets mis en avant dans l'actualité.



Recommandation 11

Sensibiliser les réutilisateurs potentiels à l'intérêt de l'open data en communiquant et en montrant des exemples de réutilisation quand ils existent. Adapter la sensibilisation au public cible.

Dans la phase de **maintenance**, l'entité continue à publier des jeux de données, et met à jour ceux qui sont publiés. Il est crucial de publier les versions les plus mises à jour possible des jeux de données. En effet, les services développés sur base de ces données doivent pouvoir donner l'information la plus correcte possible. Il est recommandé de fournir une API pour accéder aux jeux de données changeant fréquemment. Cela permettra aux services développés sur base de ces données de récupérer la version mise à jour automatiquement. De plus, une absence de données mises à jour et longitudinales, couvrant donc plusieurs intervalles de temps successifs, est un frein important à la réutilisation (Crusoe et al., 2019), car des données sur plusieurs périodes sont nécessaires pour construire des modèles prédictifs fiables. Il est dès lors important que les utilisateurs puissent aussi accéder aux données concernant les périodes passées.



Recommandation 12

Fournir des données mises à jour et longitudinales quand c'est possible via des API, et permettre aux utilisateurs d'accéder aux données concernant les périodes passées.



PROTOTYPE

Prédiction du stationnement

Dans le cadre du cours de projet en Data Analytics, des étudiants de l'Université de Namur, des étudiants ont entrepris de prédire grâce à un modèle de machine learning l'occupation des places de stationnement en ville. Sur le portail ODWB, on retrouve des données utiles portant sur le stationnement dans plusieurs villes, par exemple les données mises à jour de l'occupation des places de Shop&Drive et Kiss&Ride à Liège.

Adresse	Nom	Statut	Catégorie	Type	Localisation	Date du statut	Uri
1	VA04	Libre	SHOPANDDRIVE	ParkingSpot	50.61347814, 5.58859923	30 July 2024 20:32	https://www.liege.be/shop
2	PM02	Occupé	SHOPANDDRIVE	ParkingSpot	50.63868486, 5.57411481	30 July 2024 20:32	https://www.liege.be/shop
3	SL10	Libre	SHOPANDDRIVE	ParkingSpot	50.65030857, 5.59377509	30 July 2024 20:32	https://www.liege.be/shop
4	MI02	Occupé	SHOPANDDRIVE	ParkingSpot	50.64560555, 5.56633357	30 July 2024 20:31	https://www.liege.be/shop
5	BT09	Libre	SHOPANDDRIVE	ParkingSpot	50.619286, 5.56379269	30 July 2024 20:31	https://www.liege.be/shop
6	FT02	Occupé	KISSANDRIDE	ParkingSpot	50.62681424, 5.58200159	30 July 2024 20:31	https://www.liege.be/kissa
7	SL09	Occupé	SHOPANDDRIVE	ParkingSpot	50.6502721, 5.59366006	30 July 2024 20:30	https://www.liege.be/shop
8	BU03	Occupé	SHOPANDDRIVE	ParkingSpot	50.62191648, 5.57233523	30 July 2024 20:30	https://www.liege.be/shop
9	WA06	Libre	KISSANDRIDE	ParkingSpot	50.62983438, 5.55673387	30 July 2024 20:30	https://www.liege.be/kissa
10	BT06	Occupé	SHOPANDDRIVE	ParkingSpot	50.61932684, 5.56393351	30 July 2024 20:29	https://www.liege.be/shop
11	FG05	Occupé	SHOPANDDRIVE	ParkingSpot	50.62675012, 5.57256186	30 July 2024 20:27	https://www.liege.be/shop
12	WA09	Occupé	SHOPANDDRIVE	ParkingSpot	50.63051634, 5.55811186	30 July 2024 20:27	https://www.liege.be/shop
13	HA02	Libre	KISSANDRIDE	ParkingSpot	50.63326436, 5.58042848	30 July 2024 20:27	https://www.liege.be/kissa
14	MI03	Occupé	SHOPANDDRIVE	ParkingSpot	50.64561238, 5.56640062	30 July 2024 20:24	https://www.liege.be/shop

Capture d'écran : Portail open data de la Ville de Liège

En disposant de chaque mise à jour précédente des données en plus de la dernière mise à jour en date, il serait possible de faire des prédictions sur l'occupation future de ces places de stationnement, pourvu qu'il y ait suffisamment de mises à jour disponibles.

Dans cette étape, il est aussi essentiel que l'entité garde un œil sur ce qui est fait avec les données, via par exemple les chiffres de la plateforme tels que le nombre de téléchargements (Bright et al., 2019). Une approche complémentaire est de dialoguer plus proactivement avec les réutilisateurs (Chokki et al., 2022b), grâce à des formulaires de feedback sur le portail ou en les identifiant et contactant. Cela permet de mieux comprendre ce qui intéresse les réutilisateurs et d'apporter des ajustements, tels que la publication de jeux de données supplémentaires concernant un thème particulièrement populaire. Cela peut aussi permettre d'identifier des erreurs dans les données et de les corriger, améliorant ainsi la qualité des données. En effet, les réutilisateurs peuvent être vus comme des "beta-testeurs" des données et remonter ainsi des problèmes de qualité qui peuvent être résolus (Nikiforova et al., 2024).



Recommandation 13

Chercher à obtenir tout feedback possible sur la réutilisation des jeux de données et ajuster l'offre et la qualité des données en fonction.



RETOUR D'XP

Evaluation d'un portail open data

Lors de l'évaluation de l'expérience utilisateur d'un portail open data menée par le NADI avec des citoyens (Simonofski et al., 2022), l'un d'entre eux a remarqué une erreur d'encodage dans un des jeux de données qu'il a consultés. Il a souhaité utiliser le formulaire de feedback pour le signaler à l'entité en charge du portail. Il est ressorti de cette évaluation que tous les citoyens n'utiliseraient pas forcément les fonctionnalités de feedback proposées mais que certains franchiront le pas tout de même, montrant l'importance des mécanismes de feedback pour avoir du retour des utilisateurs.

Une excellente source de feedback, pour les entités qui peuvent le faire, est de travailler avec établissements proposant des formations en informatique ou en data science pour créer des projets de cours impliquant d'utiliser l'open data wallon pour développer des analyses ou des services. Cela permet d'avoir un retour de la part de plusieurs dizaines d'étudiants sur la qualité des données, les données qu'ils souhaiteraient voir disponibles, et cela offre en ensemble de réutilisations qui peuvent être mises en avant avec les jeux de données. Cela a aussi l'avantage de sensibiliser une partie de la société civile à l'intérêt de l'open data, favorisant d'autres réutilisations dans le futur.

Dans le cadre de sa mission de référent académique Smart Région, le NADI met ceci en pratique et offre aux entités publiques qui le souhaitent la possibilité de proposer des jeux de données à réutiliser aux étudiants de l'Université de Namur.

Recommandation 14



Collaborer avec les établissements scolaires afin de développer des projets de cours impliquant de développer des analyses et services sur base de l'open data pour récolter des feedbacks et exemples de réutilisation des étudiants.



SERVICE

Réutilisations des données sur les parkings de la Ville de Namur

Sur le portail open data de la Ville de Namur, plusieurs réutilisations sont mises en avant. Dans le cas des emplacements de parking, la Ville de Namur peut par exemple compter sur le Système de Transport Intelligent, qui est une réutilisation en interne, et également sur une réutilisation par des étudiants, présentée sur le portail.

The screenshot shows two data reuse examples on the Open Data portal of the City of Namur. The first example is titled "UNAMUR - Etudiants DataScience - ITI PMR" and describes a project from November 2021 where students developed a tool to help people with reduced mobility find parking. The second example is titled "Ville de Namur : Système de Transport Intelligent" and describes the city's internal system for intelligent transport, which uses real-time data to improve accessibility and monitor mobility policies.

Réutilisations (2)
Cliquez plus rapide

UNAMUR - Etudiants DataScience - ITI PMR

En novembre 2021, présentation des projets développés par les étudiants de #DataScience de l'Université de Namur dans le cadre du cours #DataAnalyticsprojet. Il leur a été demandé d'appliquer les techniques d'InformationVisualization, #GraphMining et #MachineLearning vues dans d'autres cours aux sources wallonnes #OpenData.

Projet groupe 6 : ITI PMR

Public cible : Personnes à mobilité réduite souhaitant se déplacer dans Namur.

Cas d'utilisation :

1. La personne entre sa destination.
2. Une place de parking proche lui est fournie.
3. La trajectoire la plus courte qui évite les travaux de sa place de parking à sa destination finale est fournie.

Carine Libert 0 mars 2023 09:38

Ville de Namur : Système de Transport Intelligent

Afin de proposer une meilleure accessibilité multimodale et d'orienter le comportement des usagers en matière de mobilité, la Ville de Namur s'est dotée d'un Système de Transport Intelligent (STI) répondant à trois objectifs : diffuser en temps réel des informations de mobilité, promouvoir la mobilité multimodale dans un objectif de développement durable et permettre de monitorer les politiques de mobilité sur le long terme grâce aux données.

Le STI est constitué :

- de dispositifs de récolte des données (caméras de reconnaissance automatique de plaques, capteurs de qualité de l'air, ...)
- d'équipements de diffusion de l'information (panneaux à messages variables, bornes d'informations aux voyageurs TEC, ...)
- d'un tableau de bord pour les gestionnaires de la mobilité
- d'un portail citoyen qui diffuse aux usagers des informations de mobilité en temps réel (charges de trafic, temps de parcours, arrivées des bus, données sur les vélos partagés, disponibilité de stationnement, ...)

Michaël Petit 24 novembre 2020 11:19

Capture d'écran : [Portail open data de la Ville de Namur](#)

L'étape de **suppression de données** n'intervient pas toujours. Elle concerne des cas spécifiques où, par exemple, des données sensibles ou erronées ont été publiées alors qu'elles n'auraient pas dû l'être. Ce sont des situations à gérer au cas par cas, et une prudence lors de la préparation des données permet d'éviter largement ce genre de problème. Toutefois, cela peut arriver et il est recommandé d'adopter une certaine vigilance afin de détecter ces cas problématiques le plus rapidement possible. Il est recommandé d'anticiper ce genre de situation et de définir des processus et une stratégie de communication à appliquer au cas où cela se produit.



Recommandation 15

Développer en amont des processus à mettre en place pour retirer des données qui n'auraient pas dû être publiées et communiquer si nécessaire.

L'ensemble des recommandations est synthétisé ci-dessous.

RECOMMANDATIONS	
INITIATION	Recommandation 1 Au sein de l'administration, formaliser et s'accorder sur les objectifs de la publication d'open data
	Recommandation 2 Dès le début du projet de publication, chercher à sensibiliser les membres de l'administration à l'intérêt de l'open data tout en écoutant les points de vue de chacun
INVENTAIRE DES DONNÉES	Recommandation 3 Impliquer le plus grand nombre de membres de l'administration lors de l'inventaire des données
	Recommandation 4 Si l'expertise légale manque en interne, solliciter l'avis d'experts en cas de doute avant la publication des données
	Recommandation 5 Commencer par publier des jeux de données dont la préparation est moins difficile, afin de se faire la main progressivement
	Recommandation 6 Si un public cible principal pour l'open data est clairement identifié, en tenir compte pour la priorisation des jeux de données
	Recommandation 7 Si les jeux de données sont moins utilisés qu'espéré, vérifier qu'il n'y a pas d'inadéquation entre l'offre et la demande d'open data
PUBLICATION DES DONNÉES	Recommandation 8 Travailler de concert avec d'autres administrations afin que des données concernant la même chose soient publiées dans un standard commun
	Recommandation 9 Tenir compte du public cible et de ses compétences techniques lors de la mise à disposition des données
	Recommandation 10 Documenter les jeux de données avec des descriptions et autres métadonnées pour éviter que les données soient mal interprétées par les réutilisateurs
	Recommandation 11 Sensibiliser les réutilisateurs potentiels à l'intérêt de l'open data en communiquant et en montrant des exemples de réutilisation quand ils existent
MAINTENANCE	Recommandation 12 Fournir des données mises à jour et longitudinales quand c'est possible via des API, et permettre aux utilisateurs d'accéder aux données concernant les périodes passées
	Recommandation 13 Chercher à obtenir tout feedback possible sur la réutilisation des jeux de données et ajuster l'offre et la qualité des données en fonction
	Recommandation 14 Collaborer avec les établissements scolaires afin de développer des projets de cours impliquant de développer des analyses et services sur base de l'open data
SUPPL.	Recommandation 15 Développer en amont des processus à mettre en place pour retirer des données qui n'auraient pas dû être publiées et communiquer si nécessaire

Recommandations et indicateurs Smart Région

Considérant les objectifs de la Smart Région, il fait parfaitement sens de voir la publication d'open data comme un projet Smart Région. Afin de guider les porteurs de projets Smart Région dans l'évaluation de leurs projets, le Smart City Institute a développé une liste de plusieurs dizaines d'indicateurs regroupés en 12 catégories. Afin d'illustrer la portée dans le contexte de la Smart Région des recommandations proposées dans ce guide, nous avons mis en relation les 12 catégories d'indicateurs et les 15 recommandations dans la matrice suivante. Une case présente à l'intersection d'une catégorie et d'une recommandation signifie que la recommandation favorise la réalisation d'une partie des indicateurs de la catégorie.

Il est à noter que puisque ce guide concerne des projets de publication d'open data, toutes nos recommandations sont liées à la gouvernance de la donnée d'une manière ou d'une autre.

		RECOMMANDATIONS														
		INITIATION	INVENTAIRE DES DONNÉES				PUBLICATION DES DONNÉES				MAINTENANCE		SUPPR			
		ENGAGEMENT AVEC LES UTILISATEURS														
		R1 - Formaliser les objectifs	R2 - Sensibiliser en interne	R3 - Faire un inventaire des données collectif	R4 - Solliciter une expertise légale	R5 - Commencer par les données faciles	R6 - Adapter la priorisation au public	R7 - Etudier les besoins en information	R8 - Collaborer et standardiser	R9 - Adapter la mise à disposition au public	R10 - Documenter les jeux de données	R11 - Sensibiliser les réutilisateurs potentiels	R12 - Publier des données à jour et originales	R13 - Récouter du feedback des réutilisateurs	R14 - Utiliser l'open data dans des projets de cours	R15 - Définir des processus pour la suppression
INDICATEURS SMART REGION	Lien avec les besoins des citoyens															
	Soutien politique															
	Dynamique d'acteurs															
	Adhésion citoyenne															
	Leadership															
	Transversalité															
	Gouvernance de la donnée															
	Interopérabilité															
	Reproductibilité et mutualisation															
	Gestion et optimisation des processus internes															
	Innovation															
	Durabilité															

Les **besoins des citoyens**, qui figurent parmi les réutilisateurs potentiels de jeux de données au même titre que les entreprises par exemple, peuvent être pris en compte à différents moments du projet. Réfléchir aux types de citoyens qui vont réutiliser les données **(R6)** ou les consulter afin de recueillir leurs besoins en informations **(R7)** permettra d'offrir des jeux de données qui correspondent mieux à leurs besoins. De même, adapter la mise à disposition aux citoyens ciblés par la publication **(R9)** permettra de mieux répondre à leurs besoins en la matière. Enfin, de manière plus générale, être à l'écoute des attentes et difficultés des citoyens réutilisant des jeux de données **(R13)** permettra de faire davantage correspondre l'offre d'open data à leurs besoins.

Le **soutien politique** fait référence à la volonté politique de s'engager dans la publication d'open data. Formaliser les objectifs de l'open data **(R1)** est essentiel afin d'avoir un bon alignement sur les objectifs de la démarche et de formuler ces objectifs en concertation pour arriver à une proposition davantage soutenue. La sensibilisation en interne **(R2)**, et en particulier auprès des membres politiques de l'entité, peut aussi amener à gagner des soutiens.

Les **dynamiques d'acteurs** consistent à identifier les acteurs pertinents pour le projet et à les impliquer. Les recommandations proposées encouragent une identification large d'acteurs, tant en interne qu'à l'extérieur de l'entité. Formaliser les objectifs de l'open data **(R1)** amène à identifier un acteur essentiel, à savoir le réutilisateur. La sensibilisation en interne **(R2)** permet d'impliquer en informant et en gagnant du soutien auprès des collaborateurs en interne. Réaliser un inventaire collectif des données **(R3)** encourage l'implication du plus grand nombre d'acteurs dans cet exercice, et solliciter une expertise légale **(R4)** signifie impliquer un acteur externe disposant de l'expertise nécessaire. Réfléchir aux types de citoyens qui vont réutiliser les données **(R6)** ou les consulter afin de recueillir leurs besoins en informations **(R7)** requiert respectivement d'identifier et d'impliquer dans un dialogue les réutilisateurs visés. Collaborer avec d'autres entités pour standardiser les données **(R8)** encourage à penser de manière plus large dans l'implication d'acteurs, et à créer des dynamiques qui dépassent le cadre de l'entité. Adapter la mise à disposition aux citoyens ciblés par la publication requiert aussi d'identifier et de comprendre les réutilisateurs **(R9)**. La sensibilisation auprès des réutilisateurs potentiels **(R11)** et la collecte de feedback auprès des réutilisateurs **(R13)** induisent de nouveaux mécanismes d'échange avec les réutilisateurs et permet de les impliquer dans le projet de publication d'open data au sens large. Enfin, la collaboration avec les établissements scolaires **(R14)** ajoute un acteur autour de la table, disposant de ressources pour expérimenter avec les données.

Afin de favoriser l'**adhésion citoyenne** envers le projet, il est recommandé de mettre en place des processus de participation citoyenne, et donc d'impliquer les citoyens dans le projet. Cela peut se faire en dialoguant avec eux, que ce soit pour récolter leurs besoins en information (**R7**), les sensibiliser au potentiel de l'open data (**R9**), être à l'écoute de leurs feedbacks (**R13**), ou encore impliquer les citoyens étudiants dans les établissements scolaires (**R14**). Ces processus de participation permettent de mieux comprendre les attentes des citoyens et de les amener vers une perception plus favorable de l'open data.

Le **leadership** consiste à désigner des personnes responsables au sein du projet et à mener des actions de sensibilisation. Cela intervient dès la première étape du projet, où des responsables doivent coordonner les acteurs et les fédérer autour d'objectifs choisis pour l'open data (**R1**). De plus, une sensibilisation à l'intérêt de l'open data peut être menée en interne (**R2**), vers les réutilisateurs (**R9**), ou plus particulièrement vers des élèves ou étudiants (**R14**). Finalement, la mise en place de processus pour la suppression des données implique de désigner des personnes responsables pour cette partie du projet de publication.

La **transversalité** encourage à "casser les silos" et à échanger des données, des connaissances, des ressources, des pratiques, au sein et au-delà de l'entité. Cela peut par exemple s'effectuer par une sensibilisation en interne (**R2**), impliquant des échanges de points de vue sur l'open data. Impliquer le plus grand nombre dans l'inventaire de données (**R3**) est une excellente manière de casser les silos entre les services de l'entité afin de créer une base de données commune. Obtenir une expertise légale (**R4**) encourage à s'ouvrir vers l'extérieur pour aller chercher une expertise manquante. Collecter les besoins en information des citoyens (**R7**) permet aussi d'aller chercher à l'extérieur de l'entité une connaissance supplémentaire, qui sera utile pour prendre des décisions quant à la publication des données. Collaborer avec d'autres entités (**R8**) encourage l'échange de ressources, formats de données, bonnes pratiques, et d'autres. Récolter du feedback auprès des utilisateurs (**R13**) et travailler avec des élèves ou étudiants (**R14**) permet d'échanger des retours sur leur expérience avec l'offre d'open data, afin de l'améliorer.

La **gouvernance des données** concerne les outils, processus, ressources qui permettent à une entité de gérer correctement les données visées par le projet. Puisque nous parlons ici de projets de publication d'open data, toutes les recommandations (**R1** → **R15**) sont liées à la gouvernance de la donnée d'une manière ou d'une autre.

L'**interopérabilité** fait référence à la capacité d'un projet à s'interfacer avec d'autres projets. Collaborer entre entités afin de définir des standards communs **(R8)** est un aspect essentiel de cette interopérabilité, afin de faire en sorte que les jeux de données publiés soient interopérables entre eux et avec ceux d'entités différentes. Les formats dans lesquels les données sont mises à disposition **(R9)** vont aussi jouer sur l'interopérabilité. Les formats propriétaires dépendent d'outils spécifiques et sont donc moins interopérables que des formats ouverts. Enfin, la publication des données à l'aide d'API **(R12)** favorise l'interopérabilité de ces données en permettant leur intégration dans le plus grand nombre possible de services.

La **reproductibilité** et la mutualisation concerne le partage d'informations et autres mécanismes favorisant la répliquabilité du projet dans une autre entité du territoire wallon. Un élément essentiel pour parvenir à la plus grande mutualisation possible est l'échange de pratiques, connaissances, et formats de données entre entités **(R8)**. Travailler avec des établissements scolaires **(R14)** permet d'amener les élèves ou étudiants à réfléchir à ces questions de répliquabilité.

La **gestion et l'optimisation des processus internes** concernent la formalisation de rôles, de processus, et l'utilisation efficiente des ressources. La formalisation des objectifs **(R1)** donne le point de départ à la définition des processus suivants. En effet, en fonction du public visé, certains aspects de la publication vont varier. L'inventaire collectif des données **(R3)** pose un cadre collaboratif pour cette activité. Commencer à publier des jeux de données plus faciles **(R5)** permet de construire une expérience au sein de l'entité et d'utiliser les ressources plus sagement. Réfléchir aux types de citoyens qui vont réutiliser les données **(R6)** ou les consulter afin de recueillir leurs besoins en informations **(R7)** affecte définit un processus de priorisation des données à publier. La définition de processus indiquant comment documenter des jeux de données **(R10)** est aussi un aspect essentiel. Obtenir du feedback auprès des réutilisateurs **(R13)** et d'élèves ou étudiants ayant travaillé avec l'open data **(R14)** donne des points d'amélioration pour les processus existants. Enfin, définir un processus clair pour le retrait de données permet de gérer le problème de manière efficiente, s'il devait se présenter.

L'**innovation** décrit le caractère innovant du projet aux niveaux technologique, organisationnel, et collaboratif. La publication d'open data implique d'utiliser de nouveaux formats de données et outils de mise à disposition, de mettre en place de nouveaux rôles et processus, et d'initier de nouvelles collaborations avec divers acteurs. Au niveau technologique, on retrouve par exemple la mise à disposition de données dans de nouveaux formats ou présentations visuelles (**R9**), voire des API (**R12**). Au niveau organisationnel, le processus de formalisation des objectifs (**R1**) est innovant, à supposer que l'entité n'ait encore jamais publié d'open data. Il implique en effet de définir de nouveaux rôles. De même, les actions de sensibilisation en interne (**R2**) et vers les réutilisateurs potentiels (**R11**) sont des processus neufs pour l'entité. La collecte de besoins en informations (**R7**) peut impliquer l'utilisation d'approches neuves pour l'entité, telles que des focus groups ou ateliers. De plus, gérer le retrait de données publiées en open data (**R15**) est quelque chose de totalement nouveau pour l'entité s'essayant pour la première fois à la publication d'open data, ainsi que pour l'entité n'ayant jamais rencontré ce problème. Au niveau collaboratif, de nombreuses collaborations nouvelles sont possibles, que ce soit en interne pour fédérer les services autour d'un inventaire commun des données (**R3**), échanger avec d'autres entités (**R8**), obtenir du feedback auprès de réutilisateurs (**R13**), ou encore collaborer avec des établissements scolaires dans le cadre de projets de cours (**R14**). Le nombre de recommandations apportant une innovation dépend évidemment des pratiques existantes au sein de chaque entité.

La **durabilité** renvoie à l'impact du projet sur les dimensions du développement durable, que sont l'environnement, la société (impact sur les individus et collectifs), et l'économie. On observe directement que certaines recommandations impliquant un dialogue ou une collaboration avec des membres de la société, tels que les citoyens pour comprendre leur besoins (**R7**) ou les établissements scolaires pour nourrir l'enseignement avec de nouveaux projets de cours (**R14**), ont un impact sur la dimension sociétale. Cependant, de manière plus large, l'impact du projet de publication sera matérialisé par l'impact des réutilisations d'open data que la publication aura rendue possibles. Ces projets peuvent avoir un impact sur les trois dimensions du développement durable, par exemple en créant de l'emploi, de la cohésion, ou en optimisant les ressources énergétiques. Chaque recommandation vise à améliorer le processus de publication de sorte à favoriser la réutilisation des jeux de données publiés. Dès lors, toutes les recommandations (**R1** → **R15**) sont liées à la durabilité d'une manière ou d'une autre.

Message clé

Le processus de publication d'open data va affecter le potentiel de réutilisation des données. En effet, autant les développeurs que le grand public peut bénéficier de ces données. Afin de leur permettre d'en tirer le meilleur, les entités doivent comprendre comment mettre les données à disposition de la meilleure manière selon le public visé. Cela implique de penser l'offre d'open data de manière à fournir des données à jour et dans une forme aussi standardisée que possible afin de favoriser une large réutilisation. Cela implique aussi de s'engager avec les réutilisateurs potentiels pour comprendre leurs besoins et en tenir compte lors de la publication. En effet, une offre d'open data qui n'est pas en adéquation avec les besoins des réutilisateurs ne trouvera pas la demande espérée.

Route de l'innovation

Aperçu du chapitre

Ce chapitre se concentre sur la réutilisation d'open data pour créer des services innovants. Il présente six grands types de réutilisation illustrés par des exemples. Il passe ensuite en revue les étapes du processus d'innovation entreprises par les réutilisateurs et en présente les freins majeurs. Nous expliquons comment les recommandations proposées pour le processus de publication, en particulier celles concernant l'engagement avec les réutilisateurs, permettent d'alléger ces freins. Enfin, le chapitre traite plus en détails l'étape d'idéation du processus d'innovation en présentant des outils et recommandations issues des recherches menées au NADI.

Types de réutilisation

Un des grands objectifs de la publication d'open data est la réutilisation par des entrepreneurs ou entreprises pour développer des services innovants (Janssen et al., 2012 ; Attard et al., 2015). C'est d'ailleurs repris dans la Directive Open Data (Directive (EU) 2019/1024, Article 18), dont le premier critère de succès listé est "l'importance de l'augmentation de la réutilisation des documents du secteur public auxquels s'applique la présente directive, **en particulier par les PME**".

La littérature scientifique a étudié quels types de réutilisation (*business models*) pouvaient être développés à partir de l'open data (Hammel et al., 2011 ; Ferro & Osella, 2013 ; Janssen & Zuidewijk, 2014 ; Magalhaes et al., 2014 ; Zeleti et al., 2016). En particulier, les six types suivants ont été identifiés. Ces types ne sont pas mutuellement exclusifs. Une réutilisation peut en combiner les caractéristiques et correspondre à plusieurs types différents.



Les **applications à but spécifique** fournissent un service basé sur l'information qui a un seul objectif. Ces applications sont propres à un domaine tel que l'environnement, la mobilité, ou encore le tourisme. Ces réutilisations sont les plus couramment observées. En effet, elles permettent de répondre à des problématiques spécifiques et peuvent souvent être développées avec des jeux de données thématiques, dépendant donc de moins de personnes pour être publiées. Par exemple, une application orientant les utilisateurs vers des places de parking ne nécessitera que des données du service mobilité de l'entité. De plus, étant donnée la diversité des thématiques possibles, énormément d'applications différentes peuvent être développées, expliquant leur fréquence élevée.



Planificateur de balades intergénérationnelles

PROTOTYPE

Dans le cadre du cours de projet en Data Analytics, des étudiants de l'Université de Namur ont développé **Ainéide**, qui utilise les open data sur les points d'intérêts et les photos anciennes de la Ville de Namur pour proposer des itinéraires de balades personnalisables. Le service entend rassembler personnes âgées et plus jeunes autour d'une même balade, invitant les personnes âgées à se remémorer et partager des histoires grâce aux photos d'époque.



PROTOTYPE

Planificateur de trajets pour personnes à mobilité réduite

Lors du hackathon Citizens of Wallonia organisé par FuturoCité, un groupe de participants a développé le projet **Pimp My Ride**, qui est un prototype d'application permettant aux personnes à mobilité réduite d'avoir des propositions de trajets prenant en compte les aménagements PMR dont ils ont besoin pour se déplacer. Ce service permet d'avoir des trajets davantage personnalisés à l'utilisateur en évitant, par exemple, qu'une personne en fauteuil roulant se voie proposer de passer par un endroit inaccessible car dépourvu de rampe. Ce projet a remporté le prix Territoire Intelligent lors du hackathon.

 Les **applications interactives** permettent aux utilisateurs d'ajouter eux-mêmes du contenu. Dans ce cas, ce sont donc les citoyens qui remontent de la donnée au lieu d'être consommateurs de la donnée. Ce genre de service est par exemple utile pour mettre en réseau des citoyens, qui donneront des informations sur ce qu'ils recherchent ou peuvent offrir afin de se connecter avec d'autres citoyens. Cela peut également être utile pour inviter les citoyens à collecter de la donnée sur le territoire afin d'aider l'entité dans son travail.



Mise en relation des citoyens

PROTOTYPE

Lors du hackathon Citizens of Wallonia organisé par FuturoCité, un groupe de participants a développé le projet **Connecta**, qui est un prototype d'application permettant de rassembler les personnes isolées autour d'un intérêt commun : le jardinage et le respect de la nature. Les citoyens ayant ce centre d'intérêt en commun et souhaitant faire des rencontres ou ayant besoin d'aide peuvent se retrouver et échanger via ce service. Ce projet a remporté le prix Wallonie en Poche lors du hackathon.

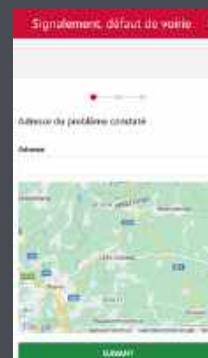


Grez-Doiseau en Poche (FixMyStreet)

SERVICE

Sur la page de la commune Grez-Doiseau sur Wallonie en Poche, les citoyens peuvent accéder au service **FixMyStreet**, qui leur permet de signaler des problèmes à l'entité, tels que des défauts de voirie. De ce fait, les citoyens peuvent faire remonter des données. Ce type d'application peut être utile dans l'étape d'inventaire des données, si les données n'existent pas encore.

Capture d'écran : [Site Wallonie en Poche de Grez-Doiseau](#)



 Les **agrégateurs d'information** combinent et traitent plusieurs sources d'open data pour les présenter aux utilisateurs. Ces services favorisent le croisement de données et augmente donc fortement le potentiel de l'open data, puisque les jeux de données ne sont plus considérés comme des données isolées, mais sont enrichis ensemble pour proposer un service. Toutefois, ces services sont souvent difficiles à mettre sur pied pour les développeurs à cause du manque de standardisation entre les données (Crusoe et al., 2019), ce qui implique beaucoup de traitement de données supplémentaires de leur côté. Cela peut expliquer pourquoi les services de ce type sont moins fréquents.



Emplacement d'un futur commerce

PROTOTYPE

Dans le cadre du cours de projet en Data Analytics, des étudiants de l'Université de Namur ont développé **EasyFind**, qui est un service permettant aux commerçants de déterminer l'emplacement idéal pour l'ouverture de leur commerce. L'application utilise plusieurs sources de données telles que le nombre de commerces déjà présents, les données démographiques, ou encore la présence de certains établissements (par exemple, les écoles). En croisant ces informations, le service est capable de recommander une zone potentielle d'implantation.



Les **modèles de comparaison** agrègent des données issues de plusieurs entités publiques afin de faire des comparaisons. Tout comme les agrégateurs de données, ce type de service étend le potentiel de l'open data en croisant des données qui ne dépassent pas l'échelle de chaque territoire. Ce type de service démontre encore une fois l'intérêt de penser à une échelle plus large que le territoire quand on publie de l'open data, puisque les réutilisations ne se limitent pas qu'au territoire.



Portails agrégateurs de données

SERVICE

En plus des plateformes spécifiques à certaines entités (par exemple, la Ville de Namur), il existe des portails open data qui, à un plus haut niveau, regroupent les données de plusieurs entités. C'est le cas par exemple du **portail open data ODWB** ou encore du **portail open data national data.gov**. Sur ce portail, on peut ainsi retrouver des jeux de données sur les subsides attribués par différentes communes.

The screenshot shows two data portals on the data.gov.be website. The left portal is titled 'Subsides octroyés en 2023' and provides an inventory of subsidies granted by the City of Brussels in 2023. The right portal is titled 'Subsides attribués par la Ville de Namur' and lists subsidies granted by the City of Namur. Both portals offer download options for various data formats: Public Sector, CSV, Brussels (Region), TTL, XML, XLSX, RDF/XML, HTML, JSON, and JSONLD. The right portal also has a 'Wallonia' button.

Capture d'écran : Portail open data data.gov.be

Bien que ce portail ne propose pas de fonctionnalités supportant explicitement la comparaison, il agit bel et bien indirectement comme un modèle de comparaison en permettant aux utilisateurs d'obtenir les données de plusieurs entités via la même plateforme. Toutefois, il revient aux utilisateurs de faire la comparaison eux-mêmes. Si les données ne sont pas standardisées dans le même format, il est difficile pour un développeur de les intégrer dans un autre outil car cela va nécessiter du traitement supplémentaire. Dans le cas des données de subsides, le montant est indiqué dans une colonne appelée *Budget Final* pour la Ville de Namur et *montant_octroye_toegekend_bedrag* pour la Ville de Bruxelles. Cette différence de nom doit être traitée manuellement dans le code du service qui sera développé à partir de ces données.

 Les **dépôts d'open data** rendent les jeux de données disponibles et accessibles pour les utilisateurs. Ce type correspond essentiellement aux plateformes publiant les données telles que les portails open data. De cette manière, ce type est un prérequis à la réalisation des autres types de réutilisation. Il est toutefois important de nuancer et de bien comprendre cette notion d'utilisateur. En effet, un utilisateur n'est pas l'autre et les portails open data doivent pouvoir rendre les données accessibles d'une manière efficace en fonction du profil des utilisateurs. Comme nous l'avons vu au travers de ce guide, le grand public et les développeurs de services, par exemple, n'ont pas le même profil en termes d'objectifs et d'expertise technique (Simonofski et al., 2022). Les données doivent donc leur être rendues disponibles de manière différente, en fonction des utilisateurs que le portail souhaite cibler.



PROTOTYPE



RETOUR D'XP

Gamification d'un portail open data

Dans ses recherches, le NADI a rapidement observé que le grand public et les développeurs de services sont deux publics très différents qui, même s'ils partagent des attentes communes envers les portails open data, ont aussi des attentes spécifiques. Il est apparu que les portails open data actuels correspondaient davantage aux développeurs (présence de formats lisibles par ordinateur, d'API, de documentation technique, etc.) mais beaucoup moins au grand public, qui pouvait se retrouver dépassé en explorant ces portails. Le NADI a exploré le potentiel de la **gamification pour rendre l'exploration d'open data plus attractive et accessible au grand public** (Simonofski et al., 2022). Le résultat est le développement d'un prototype de portail open data modifié à partir du portail existant de la Ville de Namur, avec certains ajouts tels qu'un quiz, un espace de questions d'autres citoyens, et une exploration plus visuelle des catégories de jeux de données. Ce prototype a été testé avec plusieurs citoyens qui ont confirmé que ce type de portail correspondait mieux à leurs attentes.



Cette recherche du NADI montre l'importance de l'engagement avec les utilisateurs d'open data, à plusieurs moments du projet, tant pour comprendre quelles sont leurs attentes que pour valider les solutions développées. L'engagement avec les utilisateurs ne doit pas nécessairement prendre une trop grande ampleur si les ressources sont limitées. La littérature scientifique a montré qu'**impliquer 5 utilisateurs permettait déjà d'obtenir un riche feedback** (Nielsen & Molich, 1990). Il vaut mieux impliquer 5 utilisateurs à plusieurs moments que davantage d'utilisateurs à un seul moment.

 Les **plateformes de service** offrent des fonctionnalités telles que le filtrage et des visualisations pour aider les utilisateurs à travailler avec les données. Ces plateformes vont donc un pas plus loin que donner accès aux données via des API ou des téléchargements. De nombreux portails open data ont aujourd'hui franchi ce stade et proposent de telles fonctionnalités. Une tendance émergente faisant suite au succès de ChatGPT est le développement d'un chatbot facilitant l'exploration de données, en redirigeant l'utilisateur vers des sources correspondant à sa question et même en résumant et croisant des jeux de données automatiquement.



Open data et chatbot

PROTOTYPE

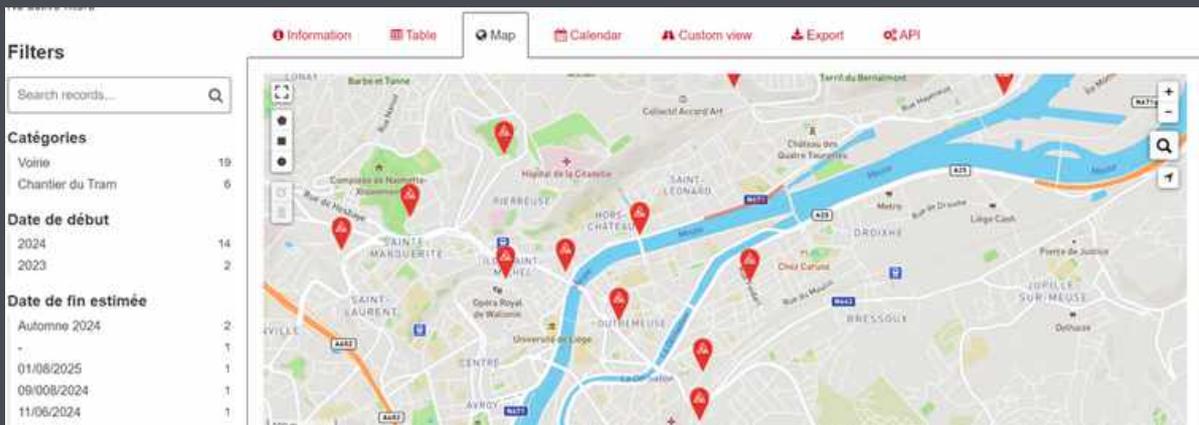
Lors du hackathon Citizens of Wallonia organisé par FuturoCité, un groupe de participants a développé le projet **DAT**, qui est un prototype d'application web permettant de formuler des requêtes en langage naturel (comme avec ChatGPT) afin d'obtenir une information se trouvant dans les jeux de données. L'utilisateur ne doit pas lui-même retrouver les jeux de données source et les traiter, le chatbot prédit la réponse automatiquement à partir des jeux de données. Toutefois, comme avec tous les systèmes d'intelligence artificielle, le taux de précision n'est pas de 100%, et il reste possible que la requête de l'utilisateur soit mal comprise et que les données ne soient pas analysées correctement. Ce projet a remporté le prix Data lors du hackathon.



Portail Open Data Wallonie-Bruxelles

SERVICE

Le portail Open Data Wallonie-Bruxelles (ODWB) (<https://www.odwb.be/pages/home/>) propose près de 1.000 jeux de données différents, issus de diverses entités de Wallonie et de Bruxelles. En plus de donner accès à ces données, le portail permet aux utilisateurs d'explorer les données à l'aide de filtres et de visualisations. Par exemple, il est possible de visualiser les travaux à Liège sur une carte et de les filtrer selon plusieurs critères tels que la catégorie et la date.



Capture d'écran : Portail Open Data Wallonie-Bruxelles



Création de tableaux de bord personnalisés

PROTOTYPE

Le NADI a affecté des recherches sur l'exploration d'open data grâce à des visualisations que l'utilisateur peut personnaliser. Les attentes des utilisateurs en ce sens ont été étudiées, et le prototype **NBDash** a été développé, permettant aux utilisateurs de choisir un jeu de données et de personnaliser des visualisations en changeant, par exemple, le type de graphique et les attributs visualisés (Chokki et al., 2022c). Les utilisateurs peuvent aussi grouper plusieurs visualisations ensemble pour constituer leur tableau de bord personnalisé.



Il y a donc une variété de types de services pouvant être développés en réutilisant de l'open data. Toutefois, la faisabilité de leur développement dépend de plusieurs facteurs. L'entité qui publie de l'open data peut, en fonction de ses ressources, en influencer certains. Par exemple, en publiant des données de bonne qualité, documentées le plus complètement possible, et structurées selon des standards communs avec d'autres entités. Ou encore, en s'engageant proactivement auprès des réutilisateurs, en les sensibilisant à l'intérêt de l'open data et en restant à l'écoute de leurs feedbacks sur l'offre d'open data. Afin de favoriser au mieux la réutilisation des open data, le processus de publication doit déjà être centré sur le réutilisateur. Pour cela, il est nécessaire de comprendre quels sont les freins que peuvent rencontrer les réutilisateurs dans le développement de services innovants basés sur l'open data. La littérature scientifique en a à plusieurs reprises fait l'inventaire. La section suivante en propose une synthèse et montre comment les entités publiques peuvent apporter une solution à ces freins en appliquant les recommandations formulées pour la Route de la publication.

Récapitulatif des types de réutilisation d'open data avec les recommandations les plus importantes pour les favoriser.

Type de réutilisation	Description	Recommandations principales
Applications à but spécifique	Résoudre un problème spécifique en donnant de l'information	Comprendre les besoins en information du public cible Publier de l'open data concernant de nombreux thèmes
Applications interactives	Permettre aux utilisateurs d'ajouter eux-mêmes du contenu / des données	Donner des incitants et rendre l'ajout de contenu facile et ludique pour motiver les utilisateurs
Agrégateurs d'informations	Combiner plusieurs jeux de données pour résoudre un problème	Adopter des standards communs lors de la publication des jeux de données
Modèles de comparaison	Comparer les données issues de plusieurs entités	Adopter des standards communs lors de la publication des jeux de données
Dépôts d'open data	Mettre à disposition des jeux de données pour les réutilisateurs	S'accorder sur les objectifs et le public cible Réaliser un inventaire des données collectif Collecter du feedback
Plateformes de service	Donner des outils (visualisations, filtres) pour travailler avec les jeux de données	Prendre en compte l'expertise technique des utilisateurs Documenter les jeux de données

Freins au développement de services basés sur l'open data

Avant d'aborder les freins au développement de services basés sur l'open data, il convient de retracer le chemin emprunté par les développeurs dans cette démarche. Il est structuré en 4 grandes parties dans la littérature scientifique (Crusoe et al., 2019), à savoir (1) la motivation derrière la réutilisation de données, (2) la recherche et l'évaluation de données, (3) l'accès aux données et leur préparation, et enfin (4) l'agrégation et la transformation des données en un service innovant.



Processus de développement de services innovants basés sur l'open data selon (Crusoe et al., 2019).

 Dans l'étape de **motivation**, le réutilisateur doit être conscient de l'existence de l'open data et avoir une raison de l'utiliser. C'est à ce moment que le réutilisateur définit ce qu'il aimerait réaliser grâce à l'open data. L'utilisation d'open data peut avoir plusieurs points de départ (Crusoe & Ahlin, 2019). Un point de départ peut être une question à laquelle on veut répondre, ou quelque chose que l'on voudrait atteindre avec l'open data, donc une réutilisation portée par une idée. Une autre possibilité est de partir de données spécifiques et de trouver quelque chose à faire avec ces données particulières. C'est une réutilisation portée par des données. Bien qu'intuitivement il semble plus sensé de partir du problème, se concentrer sur certaines données spécifiques peut avoir un intérêt, par exemple dans des projets pilotes évaluant le potentiel de certains jeux de données. Dans tous les cas, cette étape comporte une grande partie d'idéation afin de mûrir l'idée de service.

 Dans l'étape de **recherche et évaluation**, le réutilisateur se rend sur la plateforme open data où sont publiées les données afin de trouver des données prometteuses. Il utilise les fonctionnalités de la plateforme pour identifier de tels jeux de données et les évalue afin de déterminer si ces données lui permettront de réaliser ce qui a été défini à l'étape. Cette évaluation peut se faire de plusieurs manières (Crusoe et al., 2019). Une possibilité est de faire des manipulations préliminaires sur les données pour avoir un aperçu rapide de ce qu'elles peuvent permettre de réaliser. Cela peut se faire en évaluant la qualité des données, par exemple en termes de nombre de valeurs manquantes. Les métadonnées sont également utiles pour évaluer un jeu de données, en fournissant une description qui indique son contenu. Les métadonnées incluent aussi le nom des colonnes du jeu de données, qui indique ce que chaque colonne représente. Une dernière possibilité, moins répandue, est de solliciter des experts du domaine afin d'avoir un avis quant à l'exploitabilité des données.

 Lors de l'étape d'**accès et préparation**, le réutilisateur, qui a identifié des jeux de données satisfaisants, doit accéder à leur contenu. Cela peut se faire de plusieurs manières. Ils peuvent être consultés directement sur la plateforme si elle propose des fonctionnalités d'affichage de données (par exemple, dans un tableau ou une visualisation). Les fichiers de données peuvent aussi être téléchargés directement et être exploités ensuite par un programme tiers (par exemple, MS Excel) ou dans un programme informatique. Enfin, un accès automatique est aussi parfois possible si la plateforme propose une API. Ensuite, les données devront être préparées pour leur exploitation. Cette préparation, correspondant à nettoyer les données, passe par une harmonisation si plusieurs jeux de données sont impliqués, par une correction des erreurs d'encodage, et bien d'autres opérations spécifiques au cas d'utilisation. L'harmonisation consiste à modifier les jeux de données afin qu'ils partagent des standards cohérents. Par exemple, uniformiser la manière dont l'identifiant de quartier est représenté dans tous les jeux de données. Si beaucoup de jeux de données sont mobilisés et qu'il n'y a pas de standardisation en amont lors de la publication des jeux de données, ce travail devient très conséquent pour les développeurs.

 Finalement, lors de l'étape d'**agrégation et transformation**, le réutilisateur dispose de données prêtes à être exploitées les transforme afin d'en obtenir un service innovant offrant de la valeur. Cela implique souvent de combiner plusieurs jeux de données entre eux et de les exploiter dans un programme informatique.

Bien que le processus de réutilisation d'open data tel que décrit semble simple et assez direct, il est sujet à de nombreux freins. Ces freins peuvent décourager les réutilisateurs d'open data, qui décideront de pas poursuivre le processus d'innovation s'ils sont trop bloquants. Dès lors, la littérature scientifique a longuement cherché à identifier les freins les plus importants au développement de services innovants basés sur l'open data. Ce guide en propose une synthèse basée sur de nombreuses études scientifiques (Janssen et al., 2012 ; Zuiderwijk et al., 2012 ; Martin, 2013 ; Grundstrom & Lövnord, 2014 ; Martin, 2014 ; Zuiderwijk & Janssen, 2014 ; Beno et al., 2017a ; Beno et al., 2017b ; Kučera, 2017 ; Toots et al., 2017 ; Donald Shao & Saxena, 2018 ; Saxena, 2018 ; Crusoe et al., 2019 ; Ma et al., 2019 ; Ibrahim et al., 2021 ; Zuiderwijk & de Reuver, 2021). Ces études se basent sur des analyses de documents, des questionnaires, des entretiens, et des ateliers auprès de publieurs et de réutilisateurs de données. Nous présentons les freins identifiés au travers des 4 étapes du processus de réutilisation.

Motivation. Le tout premier frein à la réutilisation d'open est que peu de réutilisateurs potentiels sont au courant de son existence car peu de communication est faite autour de l'open data. Par ailleurs, certains réutilisateurs potentiels choisissent de ne pas utiliser l'open data car ils n'en voient pas l'utilité ou l'avantage ou n'ont aucun incitant à le faire. Le fait qu'ils ne voient pas l'intérêt de l'open data peut s'expliquer en partie par l'absence de communication sur les réutilisations se basant dessus, donnant une démonstration de ce qui est possible d'atteindre grâce à l'open data.

Même une fois que les réutilisateurs ont connaissance de l'existence de l'open data et voient un intérêt à l'utiliser, trouver une idée de projet à réaliser avec s'avère parfois difficile, notamment par manque de connaissance des besoins de la population, et à cause du manque de réutilisations mises en avant, faisant qu'il est plus compliqué pour les utilisateurs de trouver de l'inspiration. Pour ces raisons, les réutilisateurs rencontrent des difficultés à trouver une idée de projet qui aurait une réelle valeur ajoutée pour la population. Enfin, un autre problème qui affecte négativement la motivation de réutiliser l'open data est l'inadéquation entre l'offre et la demande de données, résultant en une faible quantité de données utiles disponibles, et s'expliquant par une méconnaissance des besoins des réutilisateurs.

Recherche et évaluation. Les difficultés rencontrées au niveau de la recherche de données sont largement spécifiques à l'interface de la plateforme sur laquelle les open data sont publiés, qui peut par exemple manquer d'un index ou de fonctionnalités de recherche suffisamment avancées, ou présenter les résultats de recherche d'une mauvaise manière.

Cependant, la littérature scientifique montre que c'est au niveau de l'évaluation des données que les freins les plus importants se trouvent. En effet, l'absence ou la faible qualité d'informations techniques sur les données ou de métadonnées (par exemple, concernant le contenu des données, la source, ou la couverture temporelle) rendent les données difficiles à réutiliser car le réutilisateur ne peut pas comprendre directement ce à quoi les données correspondent.

Un autre problème, lié à celui des métadonnées, est qu'il est difficile pour les réutilisateurs d'évaluer la qualité des jeux de données en l'absence d'informations complètes sur leur qualité. Ces manques ont pour conséquence que les données sont difficiles à comprendre, et donc à réutiliser, et que la valeur des données brutes seules n'est pas immédiatement apparente, en particulier pour les réutilisateurs moins experts. De tels manquements ont aussi pour conséquence que les réutilisateurs perdent du temps car ils doivent télécharger les données et les évaluer à partir de leur contenu.

Un autre aspect que les réutilisateurs doivent évaluer est ce qu'ils sont autorisés à faire avec les données. Cette information est fournie par la licence associée aux jeux de données, mais la complexité des licences et le nombre de conditions à remplir pour pouvoir utiliser les données sont un frein pour les réutilisateurs. Enfin, un autre frein, mentionné plus marginalement, mais pertinent dans un pays multilingue comme la Belgique, est la langue de la plateforme et des données.

Accès et préparation. L'accès aux données peut être rendu difficile par le format dans lequel les données sont publiées, par exemple parce qu'ils ne sont pas lisibles par ordinateur et donc inappropriés pour être intégrés dans des services. L'accès aux données peut aussi être entravé par la nécessité de se connecter à la plateforme, de demander une autorisation d'accès, ou de payer un certain montant. Pour les réutilisateurs privilégiant un accès automatique aux données, l'absence d'une API est un problème important. Même si une API est présente, son délai de réponse et les quotas (c'est-à-dire la limite de requêtes qu'un même utilisateur peut effectuer sur un laps de temps donné) peut poser problème également en fonction des besoins techniques du service.

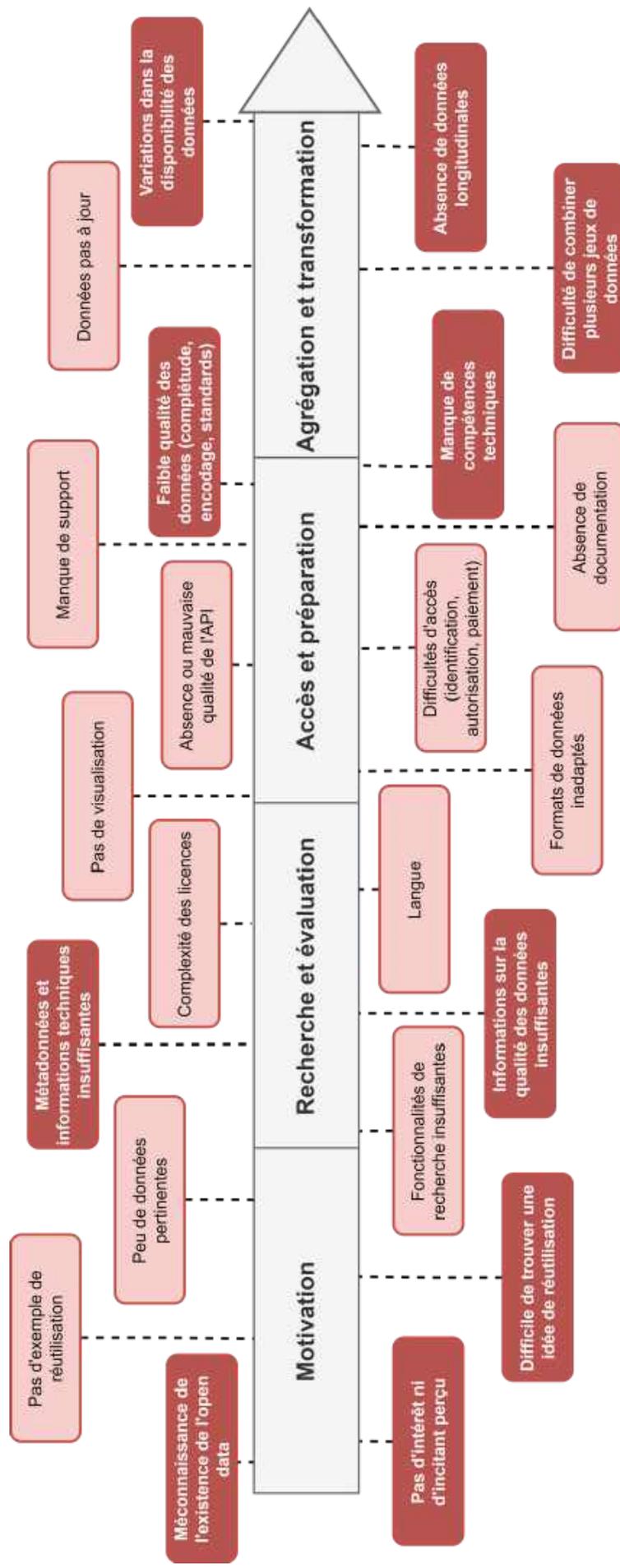
Au niveau de la préparation des données, le manque de compétences techniques devient un frein important pour les utilisateurs, même si ceux-ci ont déjà une certaine expertise, puisque c'est le moment où les manipulations techniques sur les données commencent, en préparation à leur exploitation lors de l'étape suivante.

En plus du manque de compétences, l'absence de documentation assistant les utilisateurs dans ce processus rend la tâche difficile, de même que l'absence de support de la part du publieur sur la plateforme. Les problèmes de qualité, tels que l'incomplétude des données, le manque de standards, ou encore les problèmes d'encodage, ralentissent encore la préparation des données, qui peut rapidement devenir chronophage.

Agrégation et transformation. Un frein majeur est la difficulté de combiner plusieurs jeux de données, en partie à cause du manque de standards communs au travers des jeux de données. Des exemples sont l'absence d'identifiants communs à plusieurs jeux de données, des granularités géographiques différentes (par exemple, un jeu de données par quartier et un autre par commune), une périodicité différente, ou encore des mêmes entités représentées de façon différente dans des jeux de données différents. Cependant, combiner des jeux de données est souvent nécessaire, puisqu'il arrive qu'un seul jeu de données ne donne pas une vision complète. Si les jeux de données ne peuvent pas être combinés, ce sont plusieurs types de réutilisation qui se retrouvent impossibles à réaliser.

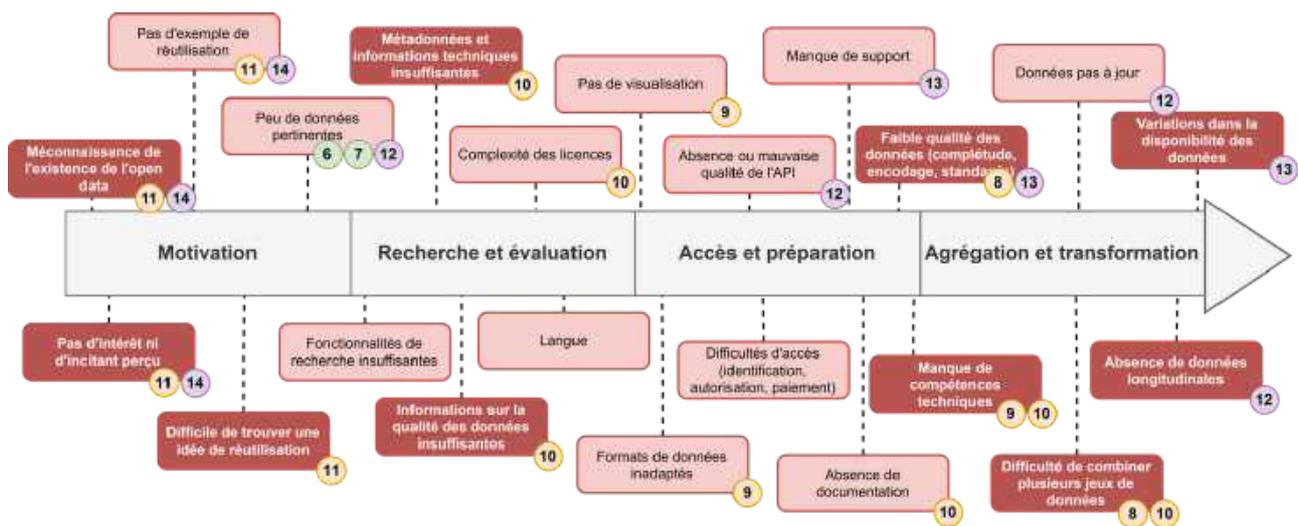
Un autre problème lié à l'exploitabilité des données est leur mise à jour . En effet, il est fréquent que les données couvrent une période passée depuis longtemps, ce qui rend leur intérêt limité. Il y a aussi un manque de données longitudinales, c'est-à-dire disposer de données mesurées de manière répétée dans le temps, nécessaires pour réaliser des analyses avancées telles que des prédictions. Enfin, un autre problème qui se présente avec des réutilisations impliquant plusieurs jeux de données est la variabilité de leur disponibilité. Si les accès aux données sont automatisés, chaque jeu de données est alors un point de défaillance potentiel, car l'indisponibilité d'un seul jeu de données met à mal tout le service.

La figure suivante reprend les freins à la réutilisation d'open data pour créer des services innovants. Les freins sont structurés selon les étapes du processus de réutilisation, et les plus importants sont indiqués en gras.



Barrières à la réutilisation d'open data dans la création de services innovants selon la littérature scientifique.

Les 15 recommandations proposées pour la Route de la publication contribuent à réduire ces freins. La sensibilisation auprès des réutilisateurs potentiels (Recommandation 11) et le développement de projets de cours (Recommandation 14) contribuent à faire connaître l'open data et son intérêt, et à avoir des exemples de réutilisation. Une bonne documentation des données (Recommandation 10) facilite non seulement la recherche de données mais peut aussi compenser un manque de compétences techniques en rendant les traitements sur les données plus simples à effectuer. Une offre d'open data dans une forme adaptée aux réutilisateurs (Recommandation 9) leur permettra d'accéder plus facilement aux données. Des efforts de standardisation dans les données (Recommandation 8) faciliteront la combinaison de jeux de données. Des données à jour et longitudinales (Recommandation 12) ouvriront la porte vers des transformations de données en services impossibles autrement et rendront donc les données plus pertinentes. Enfin, être à l'écoute du feedback des réutilisateurs (Recommandation 13) permettra d'améliorer la qualité des données et de détecter et résoudre des indisponibilités dans les données.



Lien entre les 15 recommandations formulées pour le processus de publication et les barrières à la réutilisation d'open data.

On observe que les recommandations impliquant un engagement avec les réutilisateurs (Recommandations 7, 9, 11, 13, et 14) contribuent de manière très importante à réduire les freins au développement de services innovants à partir de l'open data. Cela montre qu'une publication d'open data centrée sur les réutilisateurs, en les impliquant, est bénéfique à la réutilisation de données. Dans ce sens, l'exploitation de l'open data se joue dès le début de son processus de publication.

Trouver une idée de réutilisation

Un frein important est la difficulté de trouver une idée de réutilisation à développer. La difficulté s'explique notamment par un manque de connaissance des besoins de la population et des entités, et donc d'un problème pertinent auquel la réutilisation devrait apporter une réponse. Une autre raison est le manque d'outils à disposition des réutilisateurs pour formaliser leur idée et ainsi la raffiner en se posant les bonnes questions.

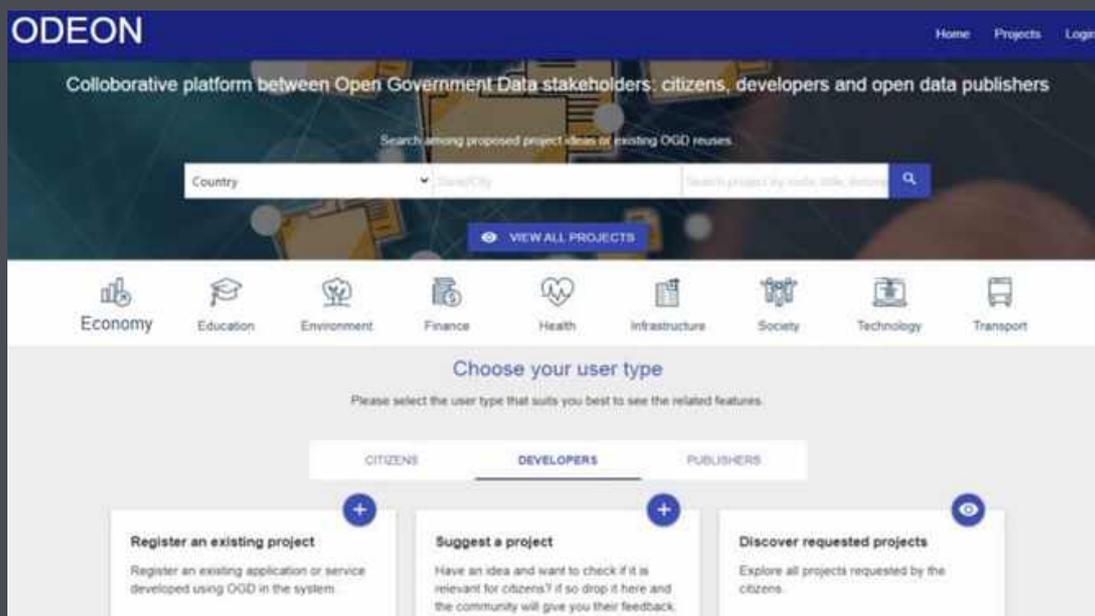
Le manque de connaissances des besoins de la population et des entités peut être logiquement comblé par des interactions entre les citoyens et/ou l'entité et les réutilisateurs, qui leur communiquent des problématiques à résoudre. Par exemple, lors du hackathon Citizens of Wallonia organisé par FuturoCité, des entités partagent des "challenges" en plus de leurs données, permettant d'orienter les réutilisateurs vers une idée à réelle valeur ajoutée. Par ailleurs, dans le cours de projet en Data Analytics donné à l'Université de Namur, les entités sont invitées à partager des problématiques et données spécifiques, orientant également la réutilisation vers un service réellement utile. Cela contribue à réduire la difficulté des réutilisateurs à trouver une idée de service innovant. Dans les deux cas, c'est à l'occasion d'un événement particulier et limité dans le temps que les entités ont remonté des problématiques vers les réutilisateurs. Cependant, cela pourrait aussi être fait de manière permanente via la plateforme open data.



PROTOTYPE

Interaction entre publieurs et réutilisateurs d'open data

Le NADI a mené une recherche débouchant sur **ODEON**, un prototype d'outil favorisant la communication entre les citoyens, les entités, et les réutilisateurs (Chokki et al., 2022b). Cet outil réunit les citoyens, entités, et réutilisateurs sur une même plateforme afin qu'ils puissent échanger des idées et des requêtes. Les citoyens peuvent consulter les services développés et en cours de développement et proposer des services qu'ils souhaiteraient voir exister. Les réutilisateurs bénéficient de cette manière d'idées précises de services qui ont une réelle utilité pour la population. Les réutilisateurs peuvent ensuite échanger avec les entités et leur demander l'accès à des jeux de données qui sont nécessaires pour développer les services proposés par la population. Le réutilisateur sert d'intermédiaire, traduisant les besoins des citoyens en besoins en données que l'entité doit satisfaire.



Recommandation 16

Davantage communiquer avec les réutilisateurs potentiels sur les besoins de la population et de l'entité afin de les aider à trouver des idées de services pertinents à développer à partir de l'open data.

Les outils permettant aux réutilisateurs de formaliser leur idée de service sont essentiels. Ces outils encouragent les réutilisateurs à se poser les bonnes questions pour avoir un regard critique sur la faisabilité du service au niveau économique, technique, ou encore légal (Arvanitis & Estevez, 2018). De plus, en mettant sur papier la proposition de valeur du service, ces outils peuvent servir de base de discussion avec d'autres personnes, comme des citoyens ou entités, afin de raffiner l'idée.

Un outil fréquemment utilisé est le Business Model Canvas (Osterwalder & Pigneur, 2010). Il se présente comme un tableau avec des cases à remplir (par exemple, proposition de valeur, sources de revenus, ressources nécessaires, clients). Il est fréquemment utilisé dans le secteur privé pour formaliser des propositions de valeur et réfléchir à la faisabilité financière d'un projet. L'utiliser pour mettre sur papier une idée de réutilisation d'open data permet au réutilisateur de réfléchir à ce que le service qu'il souhaite développer propose comme valeur ajoutée, à ce dont il a besoin pour le développer, à comment atteindre efficacement ses utilisateurs, et à comment rendre le service viable financièrement.

Cependant, les aspects portant sur les données sont trop absents du Business Model Canvas, et il est concentré exclusivement sur l'aspect financier, faisant qu'il est difficile de l'utiliser pour décrire des bénéfices sociaux et environnementaux. Par conséquent, le Business Model Canvas, bien qu'il ait l'avantage d'être répandu et simple d'utilisation, n'est pas optimal pour aider les réutilisateurs d'open data à développer leur idée de service. Une variante du Business Model Canvas a été récemment proposée dans la littérature scientifique pour tenter de palier à ses défauts (Gao & Janssen, 2022). Il décrit des aspects plus techniques tels que les fournisseurs et sources de données, ainsi que la valeur sociale de la réutilisation.

Le NADI a mené une étude comparative auprès de réutilisateurs d'open data ayant utilisé le Business Model Canvas et sa variante (Clarival et al., 2023a). Il a été observé que le Business Model Canvas est plus optimal pour traiter des aspects financiers mais sa variante est supérieure pour les aspects techniques. Sur cette base, nous proposons un Business Model Canvas étendu avec une partie spécifique aux données et une partie pour décrire l'impact de la réutilisation en termes de développement durable.

Open Data	Business Model Canvas				
Fournisseurs de données <i>D'où viennent les données ?</i>	Partenaires clés <i>Quels sont les partenaires clés nécessaires pour soutenir la proposition de valeur ?</i>	Activités clés <i>Quelles sont les activités nécessaires pour mettre en œuvre la proposition de valeur ?</i>	Proposition de valeur <i>Quelle est la valeur ajoutée apportée au client / utilisateur ? Quels problèmes la proposition de valeur résout-elle ?</i>	Relations clients <i>Quels types de relations les clients / utilisateurs s'attendent-ils à voir se développer ?</i>	Segments de clients <i>À qui s'adresse la proposition de valeur ?</i>
Intermédiaires <i>Qui traite les données brutes pour les transformer en quelque chose de plus utile, cachant la complexité des données aux utilisateurs finaux ?</i>		Ressources clés <i>Quels sont les ressources nécessaires à la mise en œuvre de la proposition de valeur ?</i>		Canaux <i>Par quels canaux de communication les clients / utilisateurs veulent-ils être atteints ?</i>	
	Structure de coûts <i>Quels coûts financiers sont inhérents à la réutilisation ?</i>		Sources de revenus <i>Pourquoi les clients / utilisateurs sont-ils prêts à payer, et comment souhaitent-ils payer ? Comment financer la réutilisation ?</i>		
Durabilité	Coûts éco-sociaux <i>Quels sont les coûts écologiques et sociaux inhérents à la réutilisation ? Quelles sont les ressources non renouvelables ?</i>		Bénéfices éco-sociaux <i>Quels sont les avantages écologiques et sociaux inhérents à la réutilisation ? Qui sont les bénéficiaires ?</i>		

Dans le cadre du hackathon Citizens of Wallonia, FuturoCité met également le CitofWal Canvas à disposition des participants. Il est utile pour formaliser une idée de réutilisation, la communiquer avec d'autres personnes, et pour servir de support visuel à un brainstorming pour raffiner l'idée. Nous ne voulons pas encourager l'utilisation d'un canevas plutôt qu'un autre dans ce guide, mais de manière plus générale encourager les entités à mettre un outil de ce type à disposition des réutilisateurs potentiels.



Plus d'informations sur le CitOfWal Canvas :

<https://www.citizensofwallonia.be/lindispensable-citofwal-canvas-2/>

Cela les aidera à développer et raffiner leur idée de service. Si les ressources le permettent, il sera judicieux de donner un feedback aux réutilisateurs ayant formulé une proposition de réutilisation, afin de les aider à construire quelque chose de plus pertinent. L'utilisation d'un canevas rend les propositions structurées de manière standard et concises, ce qui accélère l'évaluation des idées.



Recommandation 17

Mettre à disposition des réutilisateurs potentiels des outils de type canevas leur permettant de formaliser leurs idées de réutilisation et de les développer en se posant les bonnes questions.

Enfin, un outil de ce type peut aussi être complété par l'entité dans le cadre de la Recommandation 16. Les besoins en services peuvent être communiqués par le biais d'un canevas complété totalement ou partiellement.

Message clé

Les possibilités de réutilisation de l'open data sont très importantes et diverses. Cependant, afin de les réaliser, plusieurs freins que rencontrent les réutilisateurs doivent être solutionnés. Bien que l'entité n'ait pas la main sur tout, elle peut, dès la publication, être consciente de ces freins et prendre des mesures pour les alléger. Elle peut le faire en documentant correctement ses données et en les publiant dans des formats standardisés. En particulier, elle peut s'engager avec les réutilisateurs potentiels en les sensibilisant, en comprenant comment ils souhaitent disposer des données, ou encore en cherchant du feedback proactivement.

Une activité particulièrement difficile du processus de réutilisation est de trouver une idée pertinente, c'est-à-dire répondant à un besoin de la population et/ou de l'entité. L'entité peut prendre les devants et communiquer aux réutilisateurs des idées de services qui seraient pertinents à développer. Afin de rendre cette communication efficace et structurée, elle peut le faire au moyen d'outils conçus pour formaliser et communiquer des propositions de réutilisation, tels qu'un canevas.

Une publication d'open data centrée sur les réutilisateurs permettra d'alléger de nombreux freins au développement de services innovants à partir de jeux de données, et augmentera significativement le potentiel d'impact de l'open data.

Route de la participation

Aperçu du chapitre

Ce chapitre se concentre sur la réutilisation d'open data pour favoriser les projets de participation citoyenne. Il décortique le concept de participation citoyenne en distinguant trois grands types de participation. Ensuite, à partir de la littérature scientifique, il décrit les liens qui peuvent exister entre la participation citoyenne et l'open data. Cela amène à distinguer six types de projet de participation citoyenne dans lesquels l'open data peut apporter de la valeur. Le chapitre se conclut en proposant des recommandations à destination des entités afin qu'elles puissent tirer le meilleur parti possible de l'open data dans leurs projets de participation citoyenne.

La participation citoyenne et l'open data

En plus du développement de services innovants, la littérature scientifique indique que l'autre domaine de réutilisation potentielle de l'open data est la participation citoyenne (Janssen et al., 2012 ; Attard et al., 2015 ; Gonzales-Zapata & Heeks, 2015). Par participation citoyenne, il faut comprendre l'implication des citoyens, d'une manière ou d'une autre, dans les processus de prise de décision des pouvoirs publics (Callahan, 2007). Bien que ce domaine de réutilisation soit moins étudié que l'innovation dans la littérature scientifique (Kempeneer & Wolswinkel, 2013 ; Ruijter et al., 2024), ce guide souhaite accorder aux deux la même attention. En effet, la participation citoyenne est de plus en plus répandue (Royo et al., 2023), répond à une demande de la part des citoyens, et peut contribuer, en mobilisant l'intelligence collective, à trouver des solutions à des problèmes systémiques tels que le changement climatique (Namibisan & Namibisan, 2013 ; Chantillon et al., 2017). Pour cette raison, l'importance de la participation est reconnue par les Nations Unies et est un élément clé du 11e Objectif de Développement Durable (SDG 11 – Villes et communautés durables).



Plusieurs études scientifiques ont décrit des grands types de participation citoyenne, avec plus ou moins de nuance (Arnstein, 1969 ; MacIntosh, 2004 ; Ruijter et al., 2017). Globalement, la littérature scientifique s'accorde à dire qu'il y a trois grands types de participation citoyenne. On parle de **démocratie de surveillance** (liée à la notion de transparence de l'action publique), de **démocratie délibérative** dans laquelle les citoyens peuvent exprimer leur avis sur l'action publique, et enfin de **démocratie participative** où les citoyens participent encore plus activement et co-construisent l'action publique (Ruijter et al., 2017).

L'open data peut bénéficier à tous les types de participation citoyenne. En effet, pour que la participation citoyenne se déroule dans de bonnes conditions, il est nécessaire que les citoyens aient accès à suffisamment d'informations sur le sujet de la participation (Irvin, 2004). En effet, les actions de transparence nécessitent évidemment des données sur l'action publique qui soient accessibles à tous (*démocratie de surveillance*). Il en va de même quand on demande aux citoyens d'exprimer leur opinion sur l'action publique (*démocratie délibérative*). Si les citoyens ne sont pas informés correctement, cela se reflètera dans leurs opinions, qui seront mal informées (Mayeur et al., 2024). Cela peut aussi favoriser la polarisation des opinions, qui affecte négativement la qualité du débat démocratique. Enfin, quand on invite les citoyens à co-construire des solutions à un problème (*démocratie participative*), il est essentiel qu'ils disposent de suffisamment d'informations pour comprendre le problème à résoudre. Par exemple, dans le cadre de la Présidence belge du Conseil de l'Union Européenne (2024), un panel citoyen a été organisé afin de comprendre le point de vue des citoyens belges sur l'intelligence artificielle et de les inviter à formuler des propositions. Les participants ont bénéficié de sessions d'information avant de commencer les discussions. La Convention Citoyenne pour le Climat en France (2019) a également suivi cette approche.

Après avoir exposé de manière générale comment l'open data est lié à la participation citoyenne, nous élaborons plus en détail dans la section suivante la valeur que l'open data apporte aux projets de participation citoyenne.

Valeur ajoutée de l'open data dans les projets de participation citoyenne

Le NADI a mené des recherches afin de comprendre, plus précisément, dans quels cas l'open data apporte le plus de valeur ajoutée à la participation citoyenne (Clarival et al., 2024). Cela a amené à distinguer six grands types de projet dans lesquels l'open data joue un rôle utile. Ces types ont été formalisés après avoir analysé la littérature scientifique présentant des exemples de projets de participation citoyenne pouvant être liés, d'une manière ou d'une autre, à l'open data (Pak et al., 2017 ; Puussaar et al., 2018 ; Coenen et al., 2019 ; Lago et al., 2019). Ils sont décrits et illustrés avec des exemples ci-dessous.

 **Transparence passive.** Ces projets interviennent au niveau de la démocratie de surveillance. Une entité publie en open data des données décrivant l'action publique ou l'état du territoire. Un cas d'utilisation fréquent dans les projets de ce type est de communiquer des données liées aux comptes publics. Etant donné que c'est le grand public qui est visé par ce type de projet, il est recommandé de présenter les données sous forme visuelle et accessible au plus grand nombre, c'est pourquoi les portails de transparence en ligne sont une approche fréquemment utilisée. Dans ce type de projet, on ne demande pas aux citoyens de donner leur avis. On leur communique plutôt l'information et ils sont libres de manifester ou non une opinion par rapport à ce qu'ils voient. Les projets de ce type peuvent être un prérequis pour lancer des projets impliquant une participation plus active des citoyens.



Portail de transparence

SERVICE

Le **BEP (Bureau Economique de la Province)** a son propre portail open data et publie, entre autres, des données financières. Il a aussi développé un portail présentant ces données sous forme visuelle, plus accessible au grand public. Tout citoyen peut ainsi consulter les données concernant l'activité du BEP. Sur sa plateforme, le BEP met donc à disposition des données dans un objectif de transparence, mais ne demande pas explicitement aux citoyens d'utiliser ces données pour exprimer leur opinion.



Capture d'écran : [Portail du BEP](#)

Les portails de transparence ne se limitent pas à l'échelle de l'entité qui publie les données. Les données de plusieurs peuvent être agrégées afin de créer un portail plus riche, donnant une information à une échelle plus globale. Par exemple, en mettant en commun les données de qualité de l'air de nombreuses villes, il est possible de créer une cartographie à l'échelle nationale, voire européenne. Cependant, bien que combiner les données de plusieurs entités pour créer une vision plus riche augmente considérablement le potentiel de réutilisation des données, cela requiert que les données soient publiées dans un format similaire (même structure, mêmes noms de colonnes, même granularité temporelle, etc.).



SERVICE

Qualité de l'air au niveau européen

L'Agence européenne pour l'environnement met à disposition une cartographie de la qualité de l'air en temps réel à l'échelle de l'Europe. Pour fournir un tel service, elle dépend de la publication individuelle de données par des entités administratives plus locales. Cet exemple illustre que des jeux de données publiés par une entité administrative sont utiles également à d'autres niveaux de pouvoir, y compris au-delà même du pays.



Capture d'écran : [European Air Quality Index](#)



Publication des données de la participation. Ces projets sont similaires à l'idée de transparence passive mais interviennent plus tard dans le processus de participation citoyenne. L'idée ici est, à l'issue d'un projet de participation citoyenne, de publier en open data les résultats de la participation (c'est-à-dire, par exemple, le résultat d'enquêtes ou une liste d'idées proposées par les citoyens). Premièrement, cela permet de communiquer aux citoyens par rapport à ce que le projet de participation a donné. C'est essentiel que les citoyens puissent avoir accès aux résultats du processus et soient bien conscients de la manière dont ils vont être utilisés pour prendre des décisions. Sans cela, les citoyens pourraient perdre leur intérêt pour la participation citoyenne, ce qui affecterait négativement les projets futurs. Deuxièmement, cela permet à toute personne qui le souhaite de réaliser ses propres analyses sur les données de la participation.

Tout comme pour les projets de transparence passive, il est essentiel de présenter les données d'une manière accessible au plus grand nombre. Le NADI a mené des recherches sur la présentation des données de participation citoyenne (Clarival et al., 2023b) et a observé qu'une présentation visuelle peut rendre l'exploration des données plus stimulante et diminuer la surcharge d'information à laquelle font face les citoyens qui souhaitent se saisir de ces données.



SERVICE

Publication des idées des citoyens

La Ville de Liège vise élaborer son plan stratégique de manière participative avec les citoyens. Pour ce faire, la Ville a utilisé en 2017 une plateforme numérique de participation citoyenne, Réinventons Liège, à travers laquelle les citoyens peuvent proposer et voter pour des idées d'action. Ce projet de participation a débouché sur près de 1,000 idées, dont 77 ont été sélectionnées comme actions prioritaires pour la Ville. Un jeu de données reprenant ces actions a été ajouté sur la plateforme open data de la Ville.

	Catégories	Titre	Description
1	Art, Culture, Patrimoine, Tourisme	Installer des kiosques ou d'espaces ...	Renover et/ou installer des kiosques...
2	La mobilité	Créer un couloir de mobilité douce	Créer un couloir de mobilité douce ...
3	La transition énergétique	Renouveler la flotte des bus et taxis ...	Renouveler progressivement la flott...
4	La végétalisation et l'agriculture urb...	Recenser et mettre à disposition des...	Recenser et mettre à disposition les ...
5	Art, Culture, Patrimoine, Tourisme	Faire de Liège, une ville de "Street A...	Mettre l'art à portée de tous les cito...
6	Art, Culture, Patrimoine, Tourisme	Réduire l'affichage publicitaire.	La Ville de Liège a attribué la conces...
7	La mobilité	Étudier l'implantation d'un téléphéri...	Étudier l'implantation d'un téléphéri...
8	La transition énergétique	Investir dans l'énergie éolienne	Le recours aux énergies renouvelabl...
9	La végétalisation et l'agriculture urb...	Créer un marché couvert de product...	Créer un marché couvert mettant à l...
10	La végétalisation et l'agriculture urb...	Instaurer des menus bios / produits ...	Un grand nombre de repas chauds s...

Capture d'écran : [Portail Open Data de la Ville de Liège](#)



Opinion sur les données. Les projets de ce type montrent des données aux citoyens et les invite à exprimer leur opinion sur ce qu'ils voient. On parle donc de démocratie délibérative. En ce sens, ils peuvent être vus comme l'étape suivante après la transparence passive. Comme pour les projets de transparence, la présentation des données est cruciale. Elle doit être accessible et bien comprise par tous pour que chacun ait l'opportunité de s'exprimer en connaissance de cause.



Enquêtes dans l'espace public

SERVICE

Citizen Dialog Kit est une spin-off de la KULeuven qui propose une solution de participation citoyenne directement intégrée à l'espace public. Elle se présente sous la forme de petits appareils installés en rue et proposant aux passants de répondre à des questions simples en appuyant sur des boutons. Elle a par exemple été utilisée pour montrer un graphique de qualité de l'air et demander l'opinion des citoyens sur ce qu'ils voient (Coenen et al., 2019). N'importe quel jeu de donnée open data se prêtant à une représentation graphique simple (par exemple, l'évolution des subsides accordés aux projets environnementaux) pourrait être utilisé sur un dispositif de ce type pour récolter les avis des citoyens sur les données. L'exemple sur la qualité de l'air est illustré ci-dessous ainsi qu'une image schématique du dispositif Citizen Dialogue Kit.



Source : (Coenen et al., 2019)



Capture d'écran :

<https://citizendialogkit.com/>

Dans la logique d'engagement avec les utilisateurs qui est au centre de ce guide, nous recommandons aux entités mettant en œuvre ce type de projet de mener une étude pilote, avec quelques personnes, afin de vérifier que les représentations de données incluses dans le projet soient bien comprises. L'avis de 5 à 10 personnes est suffisant pour identifier la majeure partie des problèmes de compréhension potentiels.



Plus d'informations sur le nombre d'utilisateurs avec qui tester un outil : <https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/>

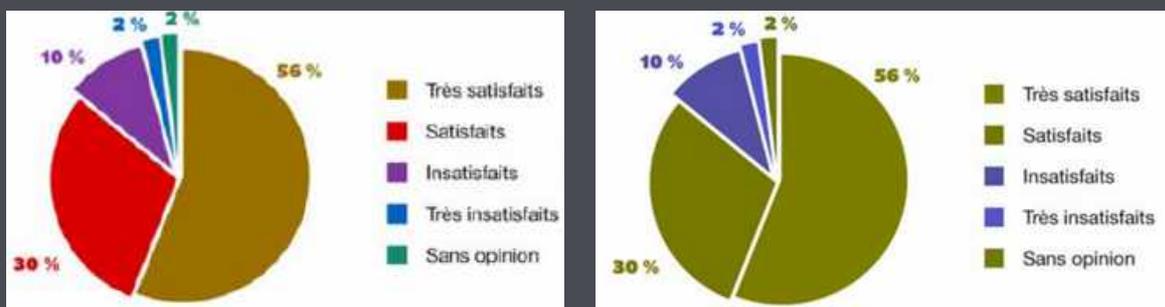
Nous recommandons aussi de porter attention à créer des visualisations aussi inclusives que possible par rapport aux handicaps des utilisateurs. Une liste de recommandations exhaustive pour une conception de représentation de données inclusives est hors du champ de ce guide, mais pour donner un exemple d'élément simple à mettre en oeuvre, le choix des couleurs sera déterminant pour les personnes souffrant de daltonisme. La visualisation suivante issue du site de Atalan (société spécialisée en accessibilité numérique en France) illustre que certaines visualisations peuvent malencontreusement exclure les utilisateurs daltoniens.



Visualisations excluant les daltoniens

PRATIQUE

Les daltoniens ne peuvent pas percevoir certaines couleurs. Cela peut les empêcher de lire correctement une visualisation de données, et donc l'exclure d'un projet de participation citoyenne se basant sur cette visualisation. L'illustration ci-dessous, créée par Atalan (société spécialisée en accessibilité numérique en France), montre comment un utilisateur daltonien (gauche) et un souffrant d'une forme de daltonisme sévère (droite) perçoivent une visualisation.



Source : <https://www.atalan.fr/agissons/fr/daltonisme.html>



Participation avec données comme support. Comme le précédent, ce type de projet cherche à recueillir l'opinion des citoyens. Il peut aussi aller plus loin en invitant les citoyens à co-construire quelque chose, dans une démarche de démocratie participative. La différence principale ici est que l'on ne demande pas aux citoyens de s'exprimer sur la donnée directement. On leur demande plutôt d'utiliser la donnée, comme support et possiblement entre d'autres choses, afin de formuler une opinion et/ou des idées les plus informées possible.



Convention Citoyenne pour le Climat

SERVICE

Organisée en 2019-2020, la Convention Citoyenne pour le Climat avait pour objectif de constituer une assemblée de citoyens français tirés au sort afin qu'ils proposent consensuellement un ensemble de mesures à mettre en place pour lutter contre le réchauffement climatique. Avant de commencer les discussions, les citoyens ont eu accès à des informations et des présentations d'experts afin de maîtriser suffisamment la problématique et de proposer des mesures aussi informées que possible.



Source :

https://www.participation-citoyenne.gouv.fr/trouver-une-concertation/83_convention-citoyenne-pour-le-climat



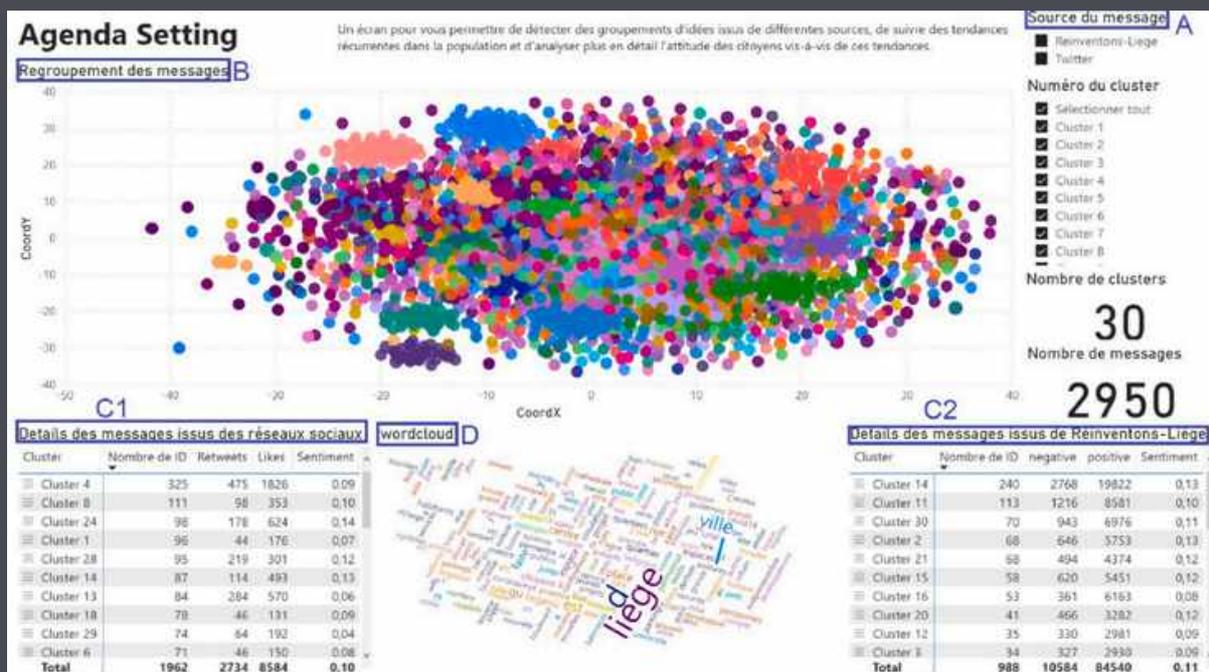
Analyse de la participation. Ce type de projet diffère des autres car il n'implique pas les citoyens dans la réutilisation d'open data. Il s'agit d'une réutilisation en interne, au sein de l'entité, de jeux de données qui vont permettre d'analyser de manière plus poussée les résultats de la participation citoyenne. Ce la va aider l'entité à prendre des décisions de plusieurs manières. La participation citoyenne peut nourrir plusieurs étapes du processus de prise de décision, en récoltant l'avis des citoyens sur le problème à résoudre en priorité, la forme qu'une solution pourrait prendre, et le choix à faire entre plusieurs possibilités de solution. Des techniques d'analyse de contenu textuel, reposant entre autres sur l'intelligence artificielle, permettent de prendre en main les données de la participation citoyenne, en produisant des résumés ou en regroupant ensemble les avis des citoyens concernant le même sujet.



PROTOTYPE

Tableau de bord d'analyse de la participation citoyenne

Lors d'une recherche menée au NADI (Simonofski et al., 2021), un prototype de tableau de bord a été créé pour illustrer comment des techniques d'intelligence artificielle peuvent aider les entités publiques à traiter les données issues de la participation citoyenne. L'outil récupère du contenu posté par les citoyens sur les réseaux sociaux et les données de la plateforme Réinventons Liège. Il déduit ensuite des catégories en regroupant ensemble les avis des citoyens grâce à une technique d'intelligence artificielle. Les résultats sont présentés aux preneurs de décision de l'entité au moyen de visualisations interactives. Ils peuvent aussi explorer des tendances dans le temps ainsi que des analyses de sentiments, montrant si les citoyens ont globalement un avis positif ou négatif sur un sujet.



De plus, utiliser des jeux de données donnant une information géographique fine, par exemple, par quartier, permet de réaliser ce type d'analyse en prenant en compte des éléments contextuels spécifiques à certaines parties du territoire. On peut de cette manière observer si une caractéristique d'une zone géographique pourrait permettre de mieux comprendre pourquoi les avis des citoyens diffèrent dans cette zone par rapport à d'autres sur un sujet donné. Par exemple, les données socio-démographiques par quartier peuvent être corrélées avec les opinions par quartier, afin d'expliquer des différences marquées qui seraient observées entre ces quartiers. Bien que les données ne doivent pas nécessairement être publiées en open data, nous recommandons qu'elles le soient, afin d'être entièrement transparent auprès des citoyens sur la manière dont les résultats de la participation sont traités. De plus, il arrive que ce soit la mise en forme de ces données en vue de les publier en open data qui permet en fait de les utiliser pour de telles analyses. Une publication large d'open data ouvre donc des perspectives intéressantes pour la prise de décisions en interne, et n'a pas un intérêt uniquement en-dehors de l'entité. Toutefois, ce type de réutilisation est plus rarement rapporté en pratique. Une explication possible est que c'est une réutilisation interne à l'entité, et donc pas nécessairement communiquée vers l'extérieur.



Analyse des opinions concernant un projet de réforme du financement de la police

PRATIQUE

Dans le projet rapporté dans l'étude scientifique de Kalampokis et al. (2011), une commune travaille sur un projet de loi visant à réduire le budget de la police. Elle recueille le contenu et les informations de localisation des messages postés par les citoyens sur les réseaux sociaux au sujet de ce projet. Ensuite, la commune combine les informations de localisation avec les autres données disponibles pour cette localisation (par exemple, la démographie de la population et le taux de criminalité) pour voir s'il existe une relation entre les opinions des citoyens et les caractéristiques de leur zone de résidentielle.



Crowdsourcing de données. Ce type de projet consiste à impliquer les citoyens dans la construction même d'un jeu de données qui sera publié en open data. En effet, s'il est possible qu'une donnée soit existante dans l'entité et qu'elle ne doive plus qu'être retravaillée pour pouvoir être publiée, il peut arriver aussi que la donnée n'existe pas. Si la donnée peut être collectée sur le territoire, l'entité peut lancer un appel aux citoyens afin qu'ils se chargent eux-mêmes, dans un effort collectif de collecter la donnée. Ce type de projet peut concerner tout type d'entité géolocalisée. Par exemple, un recensement des bancs publics avec leur géolocalisation, ou encore un recensement des panneaux routiers installés sur le territoire. La donnée ainsi collectée pourra aider l'entité dans son travail. Elle pourra aussi être réutilisée dans un futur projet de participation citoyenne ou un service innovant (par exemple, un planificateur de balades avec possibilité d'avoir des bancs où se reposer sur le trajet).



Crowdsourcing de données en Wallonie

SERVICE

FuturoCité a lancé récemment **Databusters**, une application via laquelle les entités peuvent faire appel aux citoyens pour aller « chasser la donnée » sur le territoire. Le projet a été lancé dans plusieurs communes wallonnes qui ont pu récolter des données sur l'emplacement des parkings à vélo, par exemple. L'application utilise des mécanismes de gamifications tels que des récompenses et des indicateurs de progression pour encourager les citoyens à participer.

Plus d'informations :

<https://www.futurocite.be/databusters-appli-chasseurs-donnees/>

Capture d'écran issue d'une vidéo disponible sur la chaîne YouTube de Databusters (<https://www.youtube.com/watch?v=Tyt53OuYGe4>)



Ces 6 types de projets illustrent de manière plus détaillée comment l'open data peut apporter de la valeur à la participation citoyenne, que ce soit pour informer les citoyens, leur demander leur avis, les impliquer pour construire quelque chose, mais aussi comprendre les résultats d'un projet de participation. Cependant, tous les projets de participation citoyenne sont différents, y compris au sein des types que nous avons identifiés. Il est donc tout à fait possible que parmi deux projets du même type, un tire un bien meilleur parti de l'open data que l'autre.

C'est pourquoi le NADI a étudié, avec un ensemble d'experts en open data et en participation citoyenne en Belgique et à l'international, quelles étaient les caractéristiques des projets de participation citoyenne qui tirent le plus profit de l'open data (Clarival et al., 2024). En comparant plusieurs dizaines d'exemples de projets de participation citoyenne en Wallonie mais aussi à l'international, nous avons pu déterminer que les projets qui tirent le meilleur parti de l'open data sont ceux qui :

- Concernent des thématiques spécifiques qui ont un impact dans la vie quotidienne des citoyens
- Impliquent les citoyens lors de plusieurs étapes du projet et favorisent la discussion autour des données
- Aboutissent à la production de quelque chose de concret
- Incluent des données collectées avec les citoyens
- Incluent des données géolocalisées, donc concernant des endroits précis du territoire
- Ne requièrent pas un investissement trop important de la part des citoyens

Sur cette base, nous pouvons formuler des recommandations pour chaque type de projet de participation citoyenne. Toutefois, cela ne signifie pas qu'un projet ne suivant pas ces recommandations emploie l'open data à mauvais escient, ni que le projet n'apporte pas de valeur. Nous rapportons seulement ici les éléments saillants qui se sont dégagés des recherches du NADI (Clarival et al., 2024). Ces recommandations sont à apprécier selon le contexte spécifique de chaque projet.

Nous n'incluons pas le type **Analyse de la participation** car il s'agit d'une réutilisation en interne très peu étudiée. Nous pouvons toutefois insister sur l'importance d'être transparent vis-à-vis des citoyens sur ce qui est fait avec les données de leur participation, ce qui implique de publier en open data les autres données qui ont aidé à leur analyse. Cela permettra aussi de montrer des choses concrètes auxquelles l'analyse des données de la participation a abouti.

Les projets de **transparence passive** ont l'avantage de porter sur une thématique spécifique (par exemple, l'environnement ou les budgets) et de demander un faible investissement de la part des citoyens. Par contre, la présentation des données est déterminante pour permettre au plus grand nombre d'accéder facilement aux données. De plus, des données collectées avec ou par les citoyens pourraient être intégrées afin qu'ils se sentent davantage impliqués dans une démarche qui peut sembler très unilatérale. Ces considérations s'appliquent aussi aux projets de **publication des données de la participation**.

Les projets d'**opinion sur les données** portent sur une thématique spécifique et ne demandent pas non plus trop d'investissement de la part des citoyens. Il s'agit le plus souvent de répondre à quelques questions simples. Par contre, ces projets peuvent aller plus loin que la collecte des opinions des citoyens en laissant place à une discussion sur les données. De plus, ces projets pourraient aboutir à quelque chose de concret qui serait communiqué aux citoyens, y compris par exemple un rapport sur les résultats de la participation.

Les projets de **participation avec données comme support** ont pour avantage d'encourager la discussion autour des données, ou en tout cas de la problématique décrite par les données mises à disposition. Par contre, il s'agit parfois d'appels à idées des citoyens sans thème générique. Cela peut provoquer une difficulté d'aboutir à quelque chose de concret et de naviguer au travers de toutes les données, car trop de thématiques sont abordées simultanément.

Les projets de **crowdsourcing de données** concernent une thématique spécifique, ne demandent pas trop d'investissement de la part des citoyens mais arrivent tout de même à les impliquer dans la collecte de données, et aboutissent à quelque chose de concret. Par contre, dans les cas où c'est pertinent, inclure une discussion sur les données collectées serait intéressant, et rajouterait un niveau d'engagement pour les citoyens dans le projet. Pour reprendre l'exemple des emplacements de parking vélo ou les bancs publics, le projet pourrait se poursuivre en demandant l'avis des citoyens sur les emplacements actuels et leur demander où rajouter des emplacements, si nécessaire. On observe dès lors qu'un même projet de participation citoyenne peut intégrer plusieurs types. Dans cet exemple, un projet de crowdsourcing de données se poursuit avec une phase d'opinion sur les données.

Points forts et possibilités d'amélioration des types de projets de participation citoyenne.

Type de projet	Points forts	Points d'amélioration
Transparence passive Publication des données de la participation	Thématique spécifique Pas trop d'investissement pour les citoyens	Présenter les données de manière à ce qu'elles soient accessibles au plus grand nombre Inclure des données collectées avec ou par les citoyens
Opinion sur les données	Thématique spécifique Pas trop d'investissement pour les citoyens	Inclure une discussion sur les données Viser la production de quelque chose de concret
Participation avec données comme support	Discussion autour des données	Centrer sur une problématique impactante spécifique Viser la production de quelque chose de concret
Crowdsourcing de données	Thématique spécifique Aboutit à quelque chose de concret Citoyens impliqués dans la collecte de données Pas trop d'investissement pour les citoyens	Inclure une discussion sur les données

Message clé

L'open data est réutilisé plus fréquemment dans des processus d'innovation que de participation citoyenne. Cependant, l'open data peut être bénéfique à la participation citoyenne de bien des manières. Dans une logique de démocratie de surveillance, l'open data rend compte de l'action publique et de l'état du territoire, et est donc nécessaire afin que l'information soit accessible à tous. Dans une logique de démocratie délibérative, l'open data permet aux citoyens d'avoir une base commune sur laquelle s'exprimer et formuler des suggestions. Dans une logique de démocratie participative, l'open data agit comme ressource aidant les citoyens à être informés préalablement au processus participatif. L'open data peut même être co-construit avec les citoyens, comme le montrent de nombreuses initiatives de crowdsourcing de données couronnées de succès. Dans l'ensemble, l'open data peut intervenir à différents moments de la participation citoyenne, tant préalablement comme base nécessaire à la participation, que pendant en tant que support ou objet de la participation, ou encore après dans les phases d'analyse et de communication des résultats de la participation. Nous encourageons la recherche active d'intersections entre l'open data et la participation citoyenne, encore trop rares dans la pratique.



Afin de mettre en pratique le mieux possible ces intersections entre open data et participation citoyenne, ce chapitre a mis en avant des recommandations pour chaque type de projet identifié. En particulier, premièrement, la création de représentations de données compréhensibles par le plus grand nombre, et idéalement dont la compréhensibilité est vérifiée en amont avec des citoyens. En effet, dans un processus de participation citoyenne, on vise le grand public et non plus un public de réutilisateurs experts comme c'est le cas dans les processus d'innovation. Deuxièmement, la mise en place de discussions avec les citoyens et entre les citoyens sur les données. Troisièmement, montrer aux citoyens que la participation a abouti à quelque chose de concret, que ce soit un nouveau service, des analyses, ou de nouveaux jeux de données utiles.

Conclusion

Dans ce guide, nous avons passé en revue le processus de publication d'open data et ses réutilisations possibles dans le développement de services innovants et les projets de participation citoyenne. Nous avons illustré ces processus avec de nombreux exemples et formulé des recommandations basées sur la littérature scientifique, y compris les recherches que nous menons au NADI. Pour conclure ce guide, nous souhaitons en rappeler les points les plus importants.



La publication et la réutilisation d'open data sont liées

La manière dont le processus de publication d'open data se déroule va affecter le potentiel de réutilisation des données.



L'engagement avec les réutilisateurs est essentiel

S'engager proactivement avec les réutilisateurs est nécessaire pour comprendre leurs attentes et leur donner les meilleures conditions pour réutiliser les données.



Collaborer entre entités est nécessaire

La publication d'open data doit faire l'objet de conventions communes au-delà du territoire. Standardiser les jeux de données favorise grandement à leur réutilisation.

Via ce guide, et en particulier ses recommandations pratiques et exemples inspirants, nous encourageons les entités à s'engager dans une démarche collaborative de la publication d'open data. Cette collaboration est importante avec d'autres entités pour l'échange de standards et de bonnes pratiques. Elle l'est aussi avec les réutilisateurs pour leur offrir l'open data d'une manière qui les encouragera à l'utiliser et à en retirer le meilleur.

Références scientifiques

- Arnstein, S. R. (1969). A ladder of citizen participation. *Journal of the American Institute of Planners*, 35(4), 216-224.
- Arvanitis, S., & Estevez, L. (2018). Feasibility analysis and study. In *The Emerald Handbook of Entrepreneurship in Tourism, Travel and Hospitality* (pp. 109-129). Emerald Publishing Limited.
- Attard, J., Orlandi, F., Scerri, S., & Auer, S. (2015). A systematic review of open government data initiatives. *Government Information Quarterly*, 32(4), 399-418.
- Beno, M., Figl, K., Umbrich, J., & Polleres, A. (2017). Perception of key barriers in using and publishing open data. *JeDEM-eJournal of eDemocracy and Open Government*, 9(2), 134-165.
- Beno, M., Figl, K., Umbrich, J., & Polleres, A. (2017). Open data hopes and fears: determining the barriers of open data. In *2017 Conference for E-Democracy and Open Government (CeDEM)* (pp. 69-81). IEEE.
- Böhm, C., Freitag, M., Heise, A., Lehmann, C., Mascher, A., Naumann, F., ... & Schmidt, M. (2012). GovWILD: integrating open government data for transparency. In *Proceedings of the 21st International Conference on World Wide Web* (pp. 321-324).
- Bvuma, S., & Joseph, B. K. (2019). Empowering communities and improving public services through open data: South African local government perspective. *Governance Models for Creating Public Value in Open Data Initiatives*, 141-160.
- Bright, J., Lee, S., Margetts, H., Wang, N., & Hale, S. (2019). Explaining download patterns in open government data: citizen participation or private enterprise?. *International Journal of Electronic Governance*, 11(2), 217-234.
- Callahan, K. (2007). Citizen participation: Models and methods. *International Journal of Public entité*, 30(11), 1179-1196.
- Cantador, I., Cortés-Cediel, M. E., & Fernández, M. (2020). Exploiting Open Data to analyze discussion and controversy in online citizen participation. *Information Processing & Management*, 57(5), 102301.
- Chantillon, M., Crompvoets, J., & Peristeras, V. (2017). The governance landscape of geospatial e-services—The Belgian case. *ISPRS International Journal of Geo-Information*, 6(9), 282.
-  Chokki, A. P., Simonofski, A., Frénay, B., & Vanderose, B. (2022a). Open government data awareness: eliciting citizens' requirements for application design. *Transforming Government: People, Process and Policy*, 16(4), 377-390.
-  Chokki, A. P., Simonofski, A., Clarinval, A., Frénay, B., & Vanderose, B. (2022b). Fostering interaction between open government data stakeholders: an exchange platform for citizens, developers and publishers. In *International Conference on Electronic Government* (pp. 228-243). Springer International Publishing.
-  Chokki, A. P., Simonofski, A., Frénay, B., & Vanderose, B. (2022c). Engaging citizens with open government data: The value of dashboards compared to individual visualizations. *Digital Government: Research and Practice*, 3(3), 1-20.
-  Chokki, A. P., Clarinval, A., Simonofski, A., & Vanderose, B. (2023). Evaluating a Conversational Agent for Open Government Data Quality Assessment. In *AMCIS 2023*.

-  • Clarinval, A., Simonofski, A., Castiaux, A., Gao, Y., & Janssen, M. (2023a). Formulating Open Data-Based Value Propositions: An Evaluation and Comparison of Two Canvas Tools. In *Proceedings of the 24th Annual International Conference on Digital Government Research* (pp. 327-337).
-  • Clarinval, A., Albert, J., Schelings, C., Elsen, C., Dumas, B., & Castiaux, A. (2023n). Idea Browsing on Digital Participation Platforms: A Mixed-Methods Requirements Study. In *Proceedings of the International Conference on Research Challenges in Information Science* (pp. 35-50). Springer Nature Switzerland.
- Crusoe, J. R., & Ahlin, K. (2019). Users' activities for using open government data—a process framework. *Transforming Government: People, Process and Policy*, 13(3/4), 213-236.
-  • Crusoe, J., Simonofski, A., Clarinval, A., & Gebka, E. (2019). The impact of impediments on open government data use: insights from users. In *Proceedings of the 13th International Conference on Research Challenges in Information Science* (pp. 1-12). IEEE.
-  • Crusoe, J., Simonofski, A., & Clarinval, A. (2020a). Towards a framework for open data publishers: A comparison study between sweden and belgium. In *International Conference on Electronic Government* (pp. 262-274). Springer International Publishing.
- Dawes, S. S., Vidiasova, L., & Parkhimovich, O. (2016). Planning and designing open government data programs: An ecosystem approach. *Government Information Quarterly*, 33(1), 15-27.
- Donald Shao, D., & Saxena, S. (2019). Barriers to open government data (OGD) initiative in Tanzania: Stakeholders' perspectives. *Growth and Change*, 50(1), 470-485.
- Ferro, E., & Osella, M. (2013). Eight business model archetypes for PSI re-use. In *Open data on the web workshop* (pp. 23-24). London, UK: Google Campus.
- Gagliardi, D., Schina, L., Sarcinella, M. L., Mangialardi, G., Niglia, F., & Corallo, A. (2017). Information and communication technologies and public participation: interactive maps and value added for citizens. *Government Information Quarterly*, 34(1), 153-166.
- Gao, Y., & Janssen, M. (2022). The open data canvas—Analyzing value creation from open data. *Digital Government: Research and Practice*, 3(1), 1-15.
- Gascó-Hernández, M., Martin, E. G., Reggi, L., Pyo, S., & Luna-Reyes, L. F. (2017). Citizen co-production through open data: cases of citizen training and engagement. In *Proceedings of the 18th Annual International Conference on Digital Government Research* (pp. 562-563).
- Gonzalez-Zapata, F., & Heeks, R. (2015). The multiple meanings of open government data: Understanding different stakeholders and their perspectives. *Government Information Quarterly*, 32(4), 441-452.
- Goodchild, M. F. (2007). Citizens as sensors: the world of volunteered geography. *GeoJournal*, 69, 211-221.
- Grundstrom, C., & Lövnord, A. (2014). Open Data: Barriers when adopting, implementing, and innovating open data in local municipal organizations.
- Hammell, R., Perricos, C., Branch, D., & Lewis, H. (2011). *Unlocking growth: How open data creates new opportunities for the UK*. Growth (Lakeland).
- Hung, M. J., & Hsieh, W. H. (2019). Examining the Perception and Use of Open Crime Data from a Citizen Perspective. *Chinese Public entité Review*, 10(1), 46-59.
- Hogan, M., Ojo, A., Harney, O., Ruijter, E., Meijer, A., Andriessen, J., ... & Malandrino, D. (2017). Governance, transparency and the collaborative design of open data collaboration platforms: understanding barriers, options, and needs. In *Government 3.0—Next Generation Government Technology Infrastructure and Services* (pp. 299-332). Springer, Cham.

- Hosio, S., Goncalves, J., Kostakos, V., & Riekk, J. (2014). Exploring civic engagement on public displays. *User-centric technology design for nonprofit and civic engagements*, 91-111.
- Ibrahim, A., Do Abdullah, S., & Arief, A. (2021). Benefits and Barriers of Open and One Government Data: A Systematic Review. In *IOP Conference Series: Materials Science and Engineering* (Vol. 1125, No. 1, p. 012026). IOP Publishing.
- Irvin, R. A., & Stansbury, J. (2004). Citizen participation in decision making: is it worth the effort?. *Public entité Review*, 64(1), 55-65.
- Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, adoption barriers and myths of open data and open government. *Information systems management*, 29(4), 258-268.
- Janssen, M., & Zuiderwijk, A. (2014). Infomediary business models for connecting open data providers and users. *Social Science Computer Review*, 32(5), 694-711.
- Kalampokis, E., Hausenblas, M., & Tarabanis, K. (2011). Combining social and government open data for participatory decision-making. In *Proceedings of the Third IFIP WG 8.5 International Conference* (pp. 36-47). Springer Berlin Heidelberg.
- Kempeneer, S., & Wolswinkel, J. (2023). Rethinking Open Government Data for Citizen Participation. An Introduction to a Special Issue. *Information Polity*, 28(2), 163-173.
- Kim, N. W., Jung, J., Ko, E. Y., Han, S., Lee, C. W., Kim, J., & Kim, J. (2016). Budgetmap: Engaging taxpayers in the issue-driven classification of a government budget. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* (pp. 1028-1039).
- Kucera, J., Chlapek, D., Klímek, J., & Necaský, M. (2015). Methodologies and Best Practices for Open Data Publication. In *DATESO* (pp. 52-64).
- Kučera, J. (2017). Analysis of barriers to publishing and re-use of open government data. *IDIMT 2017 Digit. Manag. Soc. Econ.-25th Interdiscip. Inf. Manag. Talks*, 305-314.
- Lago, N., Durieux, M., Pouleur, J.-A., Scoubeau, C., Elsen, C., & Schelings, C. (2019). Citizen participation through digital platforms: The challenging question of data processing for cities. *Proceedings of the 8th International Conference on Smart Cities, Systems, Devices and Technologies* (pp. 19-25).
- Ma, R., & Lam, P. T. (2019). Investigating the barriers faced by stakeholders in open data development: A study on Hong Kong as a “smart city”. *Cities*, 92, 36-46.
- Macintosh, A. (2004). Characterizing e-participation in policy-making. In *37th Annual Hawaii International Conference on System Sciences, 2004. Proceedings of the* (pp. 10-pp). IEEE.
- Magalhaes, G., Roseira, C., & Manley, L. (2014). Business models for open government data. In *Proceedings of the 8th International Conference on Theory and Practice of Electronic Governance* (pp. 365-370).
- Magalhaes, G., & Roseira, C. (2020). Open government data and the private sector: An empirical view on business models and value creation. *Government Information Quarterly*, 37(3), 101248.
- Martin, S. (2013). Risk analysis to overcome barriers to open data. *Electronic Journal of e-Government*, 11(2), pp348-359.
- Martin, C. (2014). Barriers to the open government data agenda: Taking a multi-level perspective. *Policy & Internet*, 6(3), 217-240.
- Mayeur, C., Mertes, H., & Van Hoof, W. (2024). How to Deal with Uninformed and Poorly Informed Opinions of Citizens? A Critical Approach to Online Public Engagement. *Citizen Science: Theory and Practice*, 9(1), 4.

- Nambisan, S., & Nambisan, P. (2013). Engaging citizens in co-creation in public services. IBM Center for Business Development, 8-48.
- Nikiporova, A., Clarinval, A., Zuiderwijk, A., Rudmark, D., Milic, P., & Rajamäe-Soosaar, K. (2024). Innovation Resistance Theory in Action: Unveiling Barriers to Open Government Data Adoption by Public Organizations to Unlock Open Data Innovation. arXiv preprint arXiv:2407.10883.
- Nielsen, J., & Molich, R. (1990). Heuristic evaluation of user interfaces. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 249-256).
- Noyman, A., Holtz, T., Kröger, J., Noennig, J. R., & Larson, K. (2017). Finding places: HCI platform for public participation in refugees' accommodation process. *Procedia computer science*, 112, 2463-2472.
- Osterwalder, A., & Pigneur, Y. (2010). *Business model generation: a handbook for visionaries, game changers, and challengers* (Vol. 1). John Wiley & Sons.
- Pak, B., Chua, A., & Vande Moere, A. (2017). FixMyStreet Brussels: socio-demographic inequality in crowdsourced civic participation. *Journal of Urban Technology*, 24(2), 65-87.
- Puussaar, A., Johnson, I. G., Montague, K., James, P., & Wright, P. (2018). Making open data work for civic advocacy. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW), 1-20.
- Royo, S., Bellò, B., Torres, L., & Downe, J. (2023). The success of e-participation. Learning lessons from Decide Madrid and We asked, You said, We did in Scotland. *Policy & Internet*.
- Ruijter, E., Grimmelikhuisen, S., & Meijer, A. (2017). Open data for democracy: Developing a theoretical framework for open data use. *Government Information Quarterly*, 34(1), 45-52.
- Ruijter, E., Dymanus, C., van Kesteren, E. J., Boeschoten, L., & Meijer, A. (2024). Open data work for empowered deliberative democracy: Findings from a living lab study. *Government Information Quarterly*, 41(1), 101902.
- Saxena, S. (2018). Drivers and barriers towards re-using open government data (OGD): a case study of open data initiative in Oman. *foresight*.
- Schmuderer, S., Zink, R., & Gamerith, W. (2019). Citizen participation via digital maps: A comparison of current applications. In *GI Forum-Journal of Geographic Information Science* (Vol. 7, No. 2, pp. 34-46).
- Simonofski, A., Fink, J., & Burnay, C. (2021). Supporting policy-making with social media and e-participation platforms data: A policy analytics framework. *Government Information Quarterly*, 38(3), 101590.
- Simonofski, A., Zuiderwijk, A., Clarinval, A., & Hammedi, W. (2022). Tailoring open government data portals for lay citizens: A gamification theory approach. *International Journal of Information Management*, 65, 102511.
- Sivarajah, U., Weerakkody, V., Waller, P., Lee, H., Irani, Z., Choi, Y., ... & Glikman, Y. (2016). The role of e-participation and open data in evidence-based policy decision making in local government. *Journal of Organizational Computing and Electronic Commerce*, 26(1-2), 64-79.
- Skouloudakis, A., & Christopoulou, E. (2021). Participatory budgeting: Combining smart cities and open data. In *Proceedings of the 25th Pan-Hellenic Conference on Informatics* (pp. 276-282).
- Smart City Institute (2019). *Guide pratique : Nos territoires face aux données et à leur gouvernance*.



- Toots, M., McBride, K., Kalvet, T., & Krimmer, R. (2017). Open data as enabler of public service co-creation: Exploring the drivers and barriers. In 2017 Conference for E-Democracy and Open Government (CeDEM) (pp. 102-112). IEEE.
- Ubaldi, B. (2013). Open government data: Towards empirical analysis of open government data initiatives.
- Zeleti, F. A., Ojo, A., & Curry, E. (2016). Exploring the economic value of open government data. *Government Information Quarterly*, 33(3), 535-551.
- Zuiderwijk, A., Janssen, M., Choenni, S., Meijer, R., & Alibaks, R. S. (2012). Socio-technical Impediments of Open Data. *Electronic Journal of e-Government*, 10(2), pp. 156-172.
- Zuiderwijk, A., & Janssen, M. (2014). Barriers and development directions for the publication and usage of open data: A socio-technical view. In *Open government* (pp. 115-135). Springer, New York, NY.
- Zuiderwijk, A., & de Reuver, M. (2021). Why open government data initiatives fail to achieve their objectives: categorizing and prioritizing barriers through a global survey. *Transforming Government: People, Process and Policy*.

Remerciements

Nous tenons à remercier les personnes suivantes pour leur contribution à ce guide et leur soutien.

Antoine Clarinval
Nicolas Bono Rossello
Pauline Willem
Anthony Simonofski
Hervé Jacquemin
Laetitia Ory
Isabelle Rawart

Nous remercions aussi Digital Wallonia, qui dans le cadre du Plan de Relance de la Wallonie, a soutenu financièrement l'élaboration de ce guide.



Enfin, nous vous remercions d'avoir lu ce guide et espérons qu'il vous aura inspiré. N'hésitez pas à nous contacter pour toute remarque ou question.

Contactez-nous

Prof. Anthony Simonofski
Université de Namur
Rempart la Vierge 8, 5000 NAMUR

<https://directory.unamur.be/staff/asimonof>
anthony.simonofski@unamur.be



Avec le soutien de

