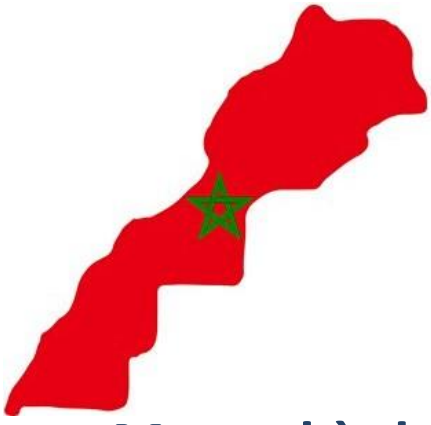


Ouverture des données publiques au Maroc

Manuel à destination des Responsables Open Data (ROD)

Avril 2021

#ADD
Agence de Développement du Digital

WORLD BANK GROUP

Table des Matières

Table des Matières 2
Introduction 3
1. Les fondamentaux de l'Open Data 4
1.1. Les concepts de l'Open Data 4
1.2. Le cadre légal de l'Open Data 8
1.3. Les acteurs de l'Open Data 9
1.4. Les bénéficiaires d'une initiative Open Data 11
2. Cycle de vie de la donnée et Approche Open Data 12
3. Rôles, fonctions et activités du ROD 16
3.1. Les interlocuteurs du ROD 16
3.2. les rôles du ROD 18
3.3. Les outils du ROD 19
3.3.1. La communication, le partage d'expérience, et l'émulation 19
3.3.2. L'inventaire de données 19
3.3.3. Les formations 20
3.3.4. Les cas d'usage 21
3.3.5. Les événements d'innovation, hackathons, et compétitions 22
3.4. Implémenter une approche Open Data 22
3.4.1. Plan annuel d'activité Open Data 23
3.4.2. Plan de suivi et évaluation 25
3.5 Les écueils à éviter 26
Annexe A – Inventaire de données 28
Annexe B – Exemple de plan de suivi et évaluation 30

Introduction :

Ce document est un manuel à destination des Responsables Open Data (ROD) désignés au sein des différentes structures publiques. L'objectif de ce manuel est d'apporter un support opérationnel au ROD et de documenter l'ensemble des tâches à mettre en œuvre pour créer une dynamique Open Data au sein de sa structure. Cette dynamique a pour objectif d'améliorer l'efficacité de l'administration, augmenter sa transparence, favoriser les collaborations entre les différentes structures publiques et avec les acteurs de la société civile et du secteur privé, et enfin favoriser l'innovation au sein de la société marocaine.

Ce manuel est divisé en trois sections. La première partie présente les concepts clés de l'Open Data, la deuxième partie présente le cycle de vie de la donnée, et enfin la troisième partie présente les fonctions et activités du Responsable Open Data.

Ce manuel a pour objectif d'être un document vivant qui s'enrichit au fur et à mesure des retours d'expérience des ROD, des défis qu'ils rencontrent et des stratégies qu'ils mettent en place pour les résoudre. Les commentaires et les propositions d'évolutions sont à envoyer à Opendata@add.gov.ma.

1. Les fondamentaux de l'Open Data :

Les données ouvertes en tant que concept trouvent leur origine dans la conviction que les données gouvernementales devraient être librement disponibles pour être publiées, utilisées et réutilisées sans limitations ni contrôle. Cette idée fondamentale s'est répandue à travers le monde depuis plus d'une décennie et accroît la demande pour des données exploitables informatiquement et libres de droits, sans limitations d'utilisation et de réutilisation. Alors que les pays développés comme le Royaume-Uni et les États-Unis ont été les premiers à s'engager dans cette voie, de plus en plus de pays en développement rejoignent aujourd'hui le mouvement et développent leur initiative nationale d'ouverture des données.

Plusieurs pays en développement ont soit lancé leur propre initiative de publication de jeux de données en ligne, soit travaillent à la mise en œuvre d'initiatives de données ouvertes dans un proche avenir. Dans le même temps, des organisations nationales et internationales, des organisations de la société civile, des institutions académiques et des entreprises du secteur privé se sont également jointes au mouvement et développent leurs capacités à exploiter ces données publiées.

Du côté de la société civile, la demande de données est généralement liée aux problématiques de la transparence et de la responsabilité des gouvernements notamment locaux et fédéraux, tandis que le secteur privé recherche de nouveaux produits et des moyens pour renforcer la valeur économique. Ce mouvement en est peut-être encore à ses balbutiements dans le contexte des pays en voie de développement mais les opportunités de développer le secteur sont immenses. Il est également utile d'explorer les points communs mais aussi les différences entre les pays développés et ceux en voie développement sur ce sujet. Les défis et les opportunités sont en effet légèrement différents. Des éléments tels que la (non) disponibilité de données sous format électronique, la collecte et la diffusion de données dans des pays ayant des siècles de tradition de secret, etc. doivent être soigneusement pris en compte. D'un autre côté, l'amélioration des services publics, l'emploi des jeunes, le développement social et économique sont des opportunités majeures.

En quelques mots, une initiative de données publiques ouvertes est une excellente occasion pour les citoyens d'un pays donné d'améliorer leur vie. Cependant, pour obtenir un tel impact, il est essentiel que les données soient publiées à temps en utilisant des formats standards et documentés pour faciliter leur réutilisation et leur exploitation par les réutilisateurs. Dans ce cadre le rôle des ROD est essentiel dans la coordination des actions au sein des différentes structures publiques.

Ce chapitre présente en détail les concepts de l'Open Data, les bases réglementaires qui gouvernent l'Open Data, les principes de la réutilisation des données publiques, les acteurs impliqués par l'ouverture des données publiques, et les rôles et fonctions des ROD.

1.1. Les concepts de l'Open Data :

Une donnée est dite « ouverte » si elle répond à deux critères :

- **La donnée est techniquement ouverte** : une donnée est techniquement ouverte si :
    - Elle est exploitable informatiquement.

- Disponible sous un format ouvert, c'est à dire non-propriétaire et documenté. Par exemple le format de fichier csv pour les données numériques. A l'inverse des formats comme les documents PDF, MS-Word (fichiers en .doc) ou MS-Excel (fichiers .xls) sont des formats propriétaires non-ouverts.

- **La donnée est légalement ouverte** : une donnée est légalement ouverte si elle est associée à des droits de réutilisation (appelés « licence de réutilisation ») large qui autorisent la réutilisation, le croisement et la redistribution de ces données, dans un contexte commercial ou pas, et requièrent uniquement la mention de la source des données. Dans le cadre de l'Open Data, le Maroc a adopté une licence spécifique présentée sur le portail Open Data[1].

A ces deux critères intrinsèques aux données, d'autres éléments sont essentiels à la réussite d'une initiative Open Data et à l'émergence des impacts attendus :

- **Les données sont publiées sous le format le plus désagrégé possible** : Le principe de l'Open Data est de publier des données dites « brutes », c'est-à-dire telles qu'elles ont été collectées au plus haut niveau de désagrégation possible afin de maximiser leur réutilisation et leur valorisation.
- **Les données brutes sont gratuites** : Un des principes fondateurs de l'Open Data est l'accès gratuit aux données. Ce principe s'appuie sur deux éléments :
  - Les données gouvernementales ont été collectées grâce aux impôts et taxes et ont donc été déjà payées par le contribuable.
  - La valorisation directe (vente des données) génère un revenu et un impact économique négligeable comparé à l'impact de la réutilisation de ces données en termes de création d'emploi, de création de richesse et de retombées directes (taxes, impôts sur la valeur ajoutée, etc.) et indirectes pour le pays.

  Il est possible que l'administration tire des revenus de ses données en offrant des services à forte valeur ajoutée sur ses propres données comme la création d'indicateurs composites ou la réalisation d'analyses spécifiques.
- **Les données doivent être facilement identifiables et accessibles** : Les données ouvertes et publiées par l'administration sont utiles et peuvent générer un impact économique et social si les réutilisateurs peuvent les trouver et y accéder facilement. Pour ce faire, la pierre angulaire d'une initiative Open Data est la mise en place d'un portail national de données ouvertes, qui est le point central d'accès à l'ensemble des données ouvertes de l'administration. Pour le cas du Maroc, le portail national a été lancé en 2011 ([2]). Ce portail héberge des jeux de données de certaines administrations, et référence des jeux de données hébergés sur des portails Open data sectoriels. Chaque structure publique a le choix, soit de mettre en place son propre

[1] http://www.data.gov.ma/fr/la-licence
[2] http://www.data.gov.ma

portail open data et le connecter au portail national pour que les jeux de données soient référencés, soit directement publier ses jeux de données sur le portail national. En termes d'accès, le téléchargement des données est libre et sans authentification afin de limiter au maximum les barrières pour les réutilisateurs.

- **Les données doivent être documentées** : Il est essentiel que les données publiées soient documentées afin de permettre leur réutilisation. La documentation des données se fait au travers de métadonnées (données sur les données). Deux types de métadonnées doivent être associées aux données :
    - **Les métadonnées descriptives**, qui correspondent à des métadonnées externes, c'est-à-dire des métadonnées typiquement utilisées pour la découverte et l'identification d'un jeu de données. Au minimum, ces métadonnées incluent le titre, la structure responsable, les coordonnées de la personne responsable au sein de la structure, la description du jeu, la couverture géographique et spatiale, la périodicité de mise à jour, le format, la licence de réutilisation et les mots clés.
    - **Les métadonnées structurelles** correspondent aux métadonnées internes, c'est-à-dire les métadonnées qui concernent la structure des objets du fichier tels que, pour un fichier de tableur par exemple, le contenu des tables et des colonnes, les clés ou les index.

  En résumé, Les métadonnées descriptives permettent la découverte de l'objet. Les métadonnées structurelles permettent d'appliquer, d'interpréter, d'analyser, de restructurer les données et de les relier à d'autres ensembles de données similaires.
- **Les données doivent être maintenues** : La qualité et donc l'impact d'un jeu de données réside, pour une partie importante, dans sa maintenance qui couvre plusieurs aspects :
    - **La mise à jour** : Les jeux de données ouvertes publiés sur le portail Open Data doivent être actualisés aussi souvent que sont mises à jour les données de références à partir desquels ils ont été construits.
    - **Les corrections et les gestions de versions :** La réutilisation d'un jeu de données et son analyse par des personnes extérieures à la structure qui en est responsable permet généralement de mettre en lumière des problèmes spécifiques. Quand ces problèmes sont remontés par les réutilisateurs, via notamment le portail national de données, il est essentiel pour le Gestionnaire de ces données de les corriger et publier des versions correctives pour s'assurer que les réutilisateurs continuent leur utilisation.
    - **Le support** : Les réutilisateurs de données sont parfois confrontés à des défis, ou ont des questions sur la structure ou le contenu. Il est essentiel pour les gestionnaires de données d'assurer le suivi et le support de leurs jeux de données pour maximiser la réutilisation. Le portail national de données ouvertes permet aux réutilisateurs de poser des questions aux gestionnaires de données, et il est donc essentiel pour ces gestionnaires d'y répondre.

- **Les données doivent être standardisées** : Enfin, la valeur intrinsèque d'un jeu de données est difficilement mesurable. L'impact émerge généralement du croisement de plusieurs jeux de données qui font émerger de nouvelles connaissances importantes. Il est donc absolument nécessaire de faciliter le travail des réutilisateurs et faciliter le croisement de jeux en adoptant des standards communs pour des données partagées entre plusieurs administrations. Ces données de références communes à l'administration sont appelées des référentiels de données et il est essentiel de s'assurer que les jeux de données ouvertes publiés les utilisent. A la date de rédaction de ce document, il est prévu qu'un ensemble de référentiels soient élaborés par l'Agence de Développement du Digital (ADD), en concertation avec le Comité de Pilotage Open Data. Il est important de noter que la mise en place et l'utilisation de référentiels est un processus à long terme. Dans un premier temps, l'administration se concentrera sur les référentiels et les standards de données les plus importants qui sont utiles à toutes les structures publiques (par exemple, le référentiel des noms géographiques ou le standard des données pour les marchés publiques). Ensuite, au fur et à mesure, des référentiels sectoriels seront adoptés pour faciliter notamment le croisement de données entre les administrations au niveau national, régional et local. Enfin, la nécessité et la volonté de concevoir et d'utiliser des référentiels et des standards de données se développeront au sein des administrations au fur et à mesure de l'augmentation de l'utilisation des données et de leur exploitation par ces administrations. Plus les équipes mettront les données au cœur de leur activité, plus elles saisiront la nécessité de ces référentiels pour faciliter leur travail, et plus elles mettront de l'énergie à les concevoir, les adopter, les diffuser et les utiliser.

L'ensemble de ces éléments sont documentés et consolidés dans la « Charte Internationale sur les données ouvertes[3] » adoptée par de nombreux pays. Cependant, ces principes couvrent presque exclusivement les méthodes de publication de données et non le contenu a proprement dit, à savoir les données concernées. Les données couvertes par l'Open Data sont définies par le cadre légal de l'Open Data qui est présenté dans la section suivante. A noter que, dans le cadre de l'Open Data, le concept de données est un concept large qui couvre tout type de données depuis les données numériques jusqu'à des textes (textes de lois par exemple) ou des données géospatiales. De la même manière, l'ouverture des données concerne une grande variété de données comme les statistiques d'un secteur, mais également des données administratives comme par exemple la liste des communes d'un département, les résultats d'une élection, les variations quotidiennes du cours d'un produit alimentaire, les arrêts et horaires de passage d'un bus ou les faits constatés par les services de police par département.

[3] https://opendatacharter.net/principles-fr/

En dehors de ces aspects techniques, un des principes clés de l'Open Data, élément de la Gouvernance Ouverte[4], est la collaboration entre l'administration et les acteurs non-gouvernementaux. L'émergence d'impacts sociaux et économiques est presque exclusivement due à la réutilisation des données ouvertes par les acteurs économiques et la société civile. Ces deux aspects, publication et réutilisation, sont très fortement liés entre eux, et l'interaction entre ces acteurs est essentielle pour la réussite d'une initiative Open Data.

1.2. Le cadre légal de l'Open Data :

L'émergence d'une initiative Open Data robuste nécessite un cadre légal clair qui définit sans ambiguïté les données qui peuvent être publiées, celles qui ne peuvent pas l'être et celles qui nécessitent un traitement avant publication. De façon générale, le développement de l'Open Data s'appuie sur deux piliers législatifs essentiels que sont les lois d'accès à l'information et les lois relatives à la protection des données personnelles.

Au Maroc, il s'agit de :

- **La Loi n°31-13 du 22 Février 2018, relative au droit d'accès à l'information**[5] : La Loi n°31-13 définit de façon précise les informations publiques qui peuvent être publiées sur demande ou de façon proactive.
- **La Loi n°09-08 du 18 Février 2009, relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel**[6] : La Loi n°09-08 fournit un cadre réglementaire pour la protection de la vie privée et identifie en particulier les données qui rentrent dans le cadre de cette loi, les traitements qui doivent être appliqués avant publication.

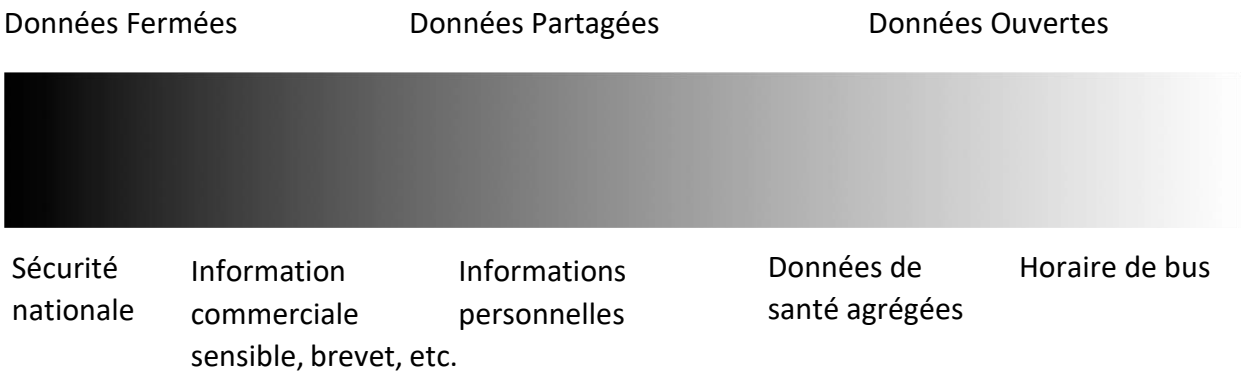
Il est important de retenir les points clés de ces législations ainsi que de la Constitution marocaine :

- L'Article 27 de la Constitution de 2011 stipule : « Les citoyennes et les citoyens ont le droit d'accéder à l'information détenue par l'administration publique, les institutions élues et les organismes investis d'une mission de service public. Le droit à l'information ne peut être limité que par la loi, dans le but d'assurer la protection de tout ce qui concerne la défense nationale, la sûreté intérieure et extérieure de l'État, ainsi que la vie privée des personnes, de prévenir l'atteinte aux droits et libertés énoncés dans la présente Constitution et de protéger des sources et des domaines expressément déterminés par la loi. »
- La Loi n°31-13 met en application cet article. Toutes les informations détenues par l'administration sont disponibles pour tout citoyen quel que soit leur format et leur type (Article 2.a) dans les limites décrites dans l'Article 7 qui interdit la publication des données suivantes :

[4] https://fr.wikipedia.org/wiki/Gouvernement_ouvert
[5] https://www.ilo.org/dyn/natlex/docs/ELECTRONIC/107094/131706/F-985611108/MAR-107094.pdf
[6] http://adala.justice.gov.ma/production/html/Fr/156780.htm

- Informations relatives de la défense nationale et à la sécurité intérieure et extérieure de l'Etat.
- Informations relatives à la vie privée.
- Informations relatives à la protection des sources d'information.
- Informations couvertes par un secret (propriété intellectuelle, brevet, secret commerciaux, délibération du Conseil des Ministres et du Conseil du Gouvernement, etc.).
- Informations impliquées dans les procédures juridiques en cours.

- Il est essentiel de noter que les lois n°31-13 et n°09-08 stipulent clairement que lorsque qu'un ensemble de données contient à la fois des données qui peuvent être publiées et des données qui ne peuvent pas être publiées, il est nécessaire de procéder au retrait des informations sensibles (anonymisation, oblitération…) tout en publiant le reste des informations publiques. Le diagramme ci-dessous montre le continuum qui existe entre les données non-publiables et les données ouvertes.



A noter qu'une analyse détaillée du cadre juridique marocain dans le cadre de l'Open Data a été réalisée en 2019-2020 par la Banque Mondiale dans le cadre de la réalisation de l'étude sur la préparation du Maroc à l'ouverture des données publiques (« ODRA »[7]).

1.3. Les acteurs de l'Open Data :

La réussite d'une initiative Open Data et la mise en place d'une dynamique positive au sein des différentes structures publiques requiert la collaboration et la coordination de plusieurs acteurs présentés ci-dessous :

- **Le Responsable de la structure** : Le responsable de la structure publique est l'acteur clé qui insuffle la vision, définit la stratégie et mobilise l'ensemble de sa structure pour la mise en œuvre d'une démarche Open Data. Sans un leadership fort qui entraine la mobilisation des personnels, l'impact de l'ouverture des données sera faible.

[7] Voir le Chapitre 2 du rapport de l'étude ODRA

- **Le Responsables Open Data (ROD)** : Le ROD est le deuxième acteur clé de la réussite de l'ouverture des données. Le ROD est l'acteur qui transforme la vision du responsable en plan d'actions, coordonne les activités, met en œuvre les différents outils et assure le suivi et évaluation afin de remonter au responsable les défis et les actions correctives à mettre en place. Le ROD a également un rôle de collaboration et coordination avec l'ADD afin de partager ses expériences, bénéficier du retour d'expérience des autres structures, et mettre en œuvre les directives globales adoptées. Le rôle et les fonctions du ROD sont détaillés dans le manuel open data à destination des Responsables Open Data[8].
- **Le Chargé d'accès à l'information** : Le chargé d'accès à l'information est un acteur indirect de l'Open Data. En tant que personne responsable de la mise en œuvre de la Loi n°31-13, il participe à l'identification des données qui peuvent ou doivent être publiées par la structure.
- **L'entité en charge des archives :** L'entité en charge des archives est un acteur indirect de l'Open Data. En tant qu'entité en charge de la mise en œuvre de la Loi n° 69.99 relative aux archives, elle participe à l'identification des documents, informations et données qui peuvent être publiées en relation avec cette loi.
- **Le Responsable des Données Personnelles (RDP) :** Le RDP est un acteur indirect de l'Open Data. En tant que personne responsable de la mise en œuvre de la Loi n°09-08, le RDP participe à l'identification des données qui relèvent de cette loi, et identifie les traitements qui doivent être appliqués aux données avant publication afin de respecter les directives de la loi.
- **Les Gestionnaires de Données (GdD)** : Les GdD sont les acteurs techniques clés de la publication de données. Ils sont en charge de la préparation et de la publication des données qu'ils gèrent et qui peuvent être publiées dans le cadre de l'ouverture des données publiques. Les GdD sont aussi en lien direct avec les réutilisateurs et interagissent avec eux sur les analyses de données ou les corrections d'erreurs potentielles. Les détails du processus de publication sont présentés dans le chapitre 3 de ce présent document.
- **Les services informatiques (SI)** : Les SI sont des acteurs au service des GdD et au travers de leur expertise informatique permettent au GdD d'automatiser la publication de données depuis la préparation jusqu'à la publication sur le portail national et/ou le portail de la structure publique concernée en passant par le transcodage et le formatage.
- **Les réutilisateurs** : Les réutilisateurs sont les acteurs clés de la valorisation des données. Par leur réutilisation, l'analyse et le croisement de données, ou l'innovation, ils permettent d'une part de générer de la valeur sociale et économique et financière à partir de la donnée. D'autre part, ils permettent également d'améliorer la qualité

[8] Réf. Manuel à destination des Gestionnaires de l'Open Data

des données en mettant à jour des incohérences potentielles. Enfin, ils aident les structures à prioriser les publications de données en fonction de leurs besoins.

1.4. Les bénéficiaires d'une initiative Open Data :

L'objectif principal d'une initiative Open Data est de générer un impact économique et social important pour le pays. Les bénéficiaires directs et indirects de l'ouverture des données sont nombreux :

- **Les structures publiques** : Les premiers bénéficiaires de l'ouverture des données publiques sont les structures publiques elles-mêmes pour deux raisons principales :
    - Elles peuvent accéder aux données des autres structures de façon simple et sans barrière administrative, ce qui facilite l'exploitation de ces données et la coordination de l'action publique entre différentes administrations. Pour illustrer ce phénomène, il est intéressant de noter par exemple, que le Gouvernement de la province de Colombie Britannique au Canada rapporte que depuis la mise en place de son portail Open Data, à peu près un tiers des requêtes de jeux de données sur ce portail viennent des structures publiques.
    - La mise en place de l'Open Data les pousse à instaurer des processus robustes de gestion de données qui, d'une part, améliore la qualité des données, et d'autres part, facilite leur exploitation pour la prise de décision et la mesure de l'efficacité des politiques publiques.
- **Les acteurs de l'innovation** : L'accès à des données publiques, notamment des données géospatiales permettent aux acteurs de l'innovation de concevoir et proposer des services innovants. Pour illustrer cette opportunité, il est intéressant de noter par exemple que NESTA[9], un organisme non-gouvernemental anglais spécialisé dans l'innovation, a investi £1.2M pour stimuler l'innovation autour des données gouvernementales du Royaume-Uni et a mesuré au bout de 3ans que cet investissement avait généré un retour sur investissement pour chaque £1 entre £5 et £10 pour l'économie anglaise soit une génération de £5.3M à £10.8M en termes de création d'emplois et de création de richesses.
- **Le secteur privé** : Les données publiques ouvertes sont une opportunité majeure pour le secteur privé, que ce soit pour améliorer l'efficacité de leurs activités comme les études de marché, et les prospections, ou pour créer de nouveaux services. A titre d'exemple, l'ODI (Open Data Institute, Institut indépendant dédié à l'Open Data[10]) a mené une étude en Angleterre qui a montré que 270 sociétés utilisaient les données publiques ouvertes, représentant un chiffre d'affaires de £92bn et plus de 500.000 emplois.

[9] https://www.nesta.org.uk/
[10] https://theodi.org/

- **La société civile** : Un des enjeux de l'Open Data est l'amélioration de la transparence de l'Administration. Grâce aux données ouvertes, les organisations de la société civile, en particulier celles qui travaillent dans ce domaine, peuvent facilement mener leurs investigations. Pour celles qui travaillent dans des secteurs spécifiques (par exemple emploi, santé…), elles peuvent trouver des données qui permettent d'orienter leurs activités.
- **Le secteur académique et universitaire** : L'ouverture des données publiques est un moteur essentiel pour la recherche universitaire. La mise à disposition de données permet aux chercheurs de conduire des recherches très précises (études sociales, économiques…) grâce aux données publiées.
- **Les médias** : L'émergence de l'Open Data dans le monde a permis également la création d'un nouveau type de journalisme, appelé Data journalisme, ou journalisme des données, où des journalistes exploitent et analyses des données ouvertes pour développer leurs articles et informer le public. De façon général, l'Open Data est un outil essentiel pour les journalistes d'investigations, et une opportunité de développer et améliorer la technicité des journalistes, et les contenus média.
- **Les citoyens** : Les citoyens sont généralement des bénéficiaires indirects de l'ouverture des données. L'Open Data leur permet d'accéder à des services innovants grâce au travail des acteurs de l'innovation, leur permet d'être mieux informer de la gestion des affaires publiques grâce aux organisations de la société civile et aux médias, et leur permet plus généralement de participer plus activement dans la vie démocratique de leur pays.

2. Cycle de vie de la donnée et Approche Open Data :

Le succès d'une initiative Open Data requiert à la fois la publication d'un nombre important de jeux de données de qualité, mis à jour régulièrement et également la réutilisation de ces données par les réutilisateurs afin de générer de la valeur économique et sociale à partir de ces données. Cependant, cette définition présente à première vue deux faiblesses importantes :

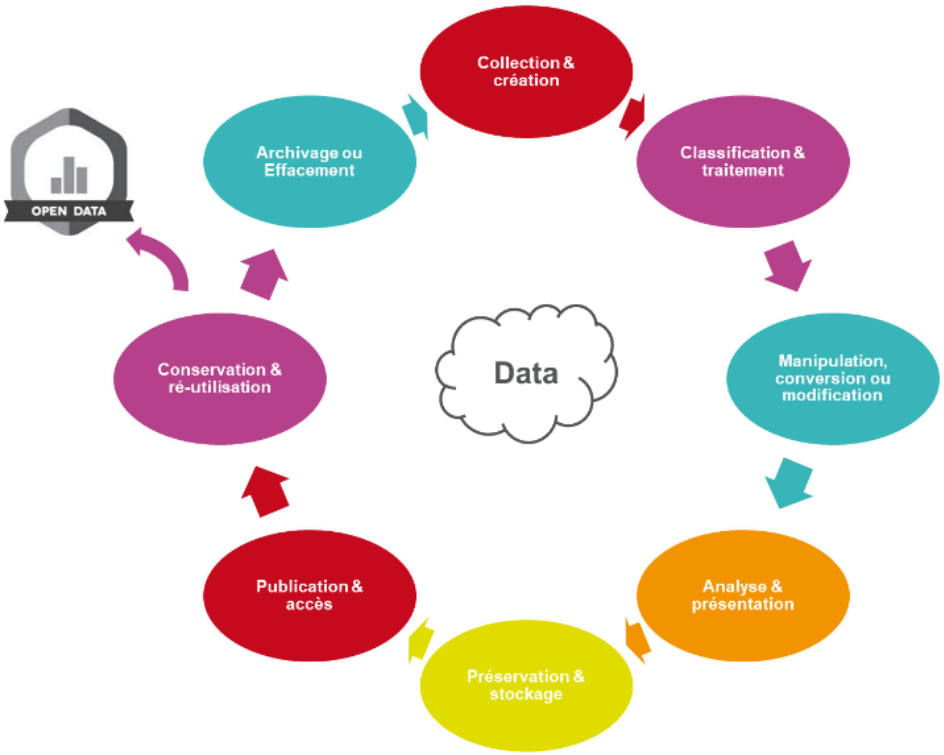
- **Une focalisation sur la publication de données** : Cette définition tend à suggérer que les structures publiques doivent principalement se focaliser sur la publication de données comme objectif et non comme le résultat d'un processus de gestion des données pérennes et robustes.
- **Un bénéfice et une génération de valeur pour les acteurs extérieurs (les réutilisateurs)** : Cette définition tend à suggérer qu'il est demandé aux structures publiques un effort et un travail (la publication de données) qui bénéficie in fine à d'autres acteurs externes à la structure (les réutilisateurs).

De nombreux pays ont tenté d'appliquer à la lettre cette définition et ils ont focalisé leur action sur la mise en ligne de données et sur la réutilisation de données publiées au travers d'activités ponctuelles comme par exemple des hackathons. Les résultats d'une telle stratégie

ont toujours été extrêmement limités et peu pérennes. D'une part, bien que la publication ponctuelle de données ouvertes soit possible, si elle n'est pas intégrée dans un processus complet de production, les données ne seront pas maintenues et mises à jour, perdant rapidement toute valeur potentielle. D'autre part, si les structures publiques qui publient des données n'y trouvent aucun bénéfice direct, leur motivation se diluera au cours du temps.

Le défi majeur de la mise en œuvre de l'Open Data est la transformation des structures publiques afin de mettre les données au centre de l'action publique. L'objectif est de mettre en place plusieurs éléments au sein des structures publiques :

- **Des processus robustes de gestion de données** où la publication de données ouvertes devient un effet de bord intégré dans la production et l'exploitation des données depuis leur collecte jusqu'à leur archivage. Le schéma ci-dessous présente la place de la publication Open Data dans le cycle de vie de la donnée.

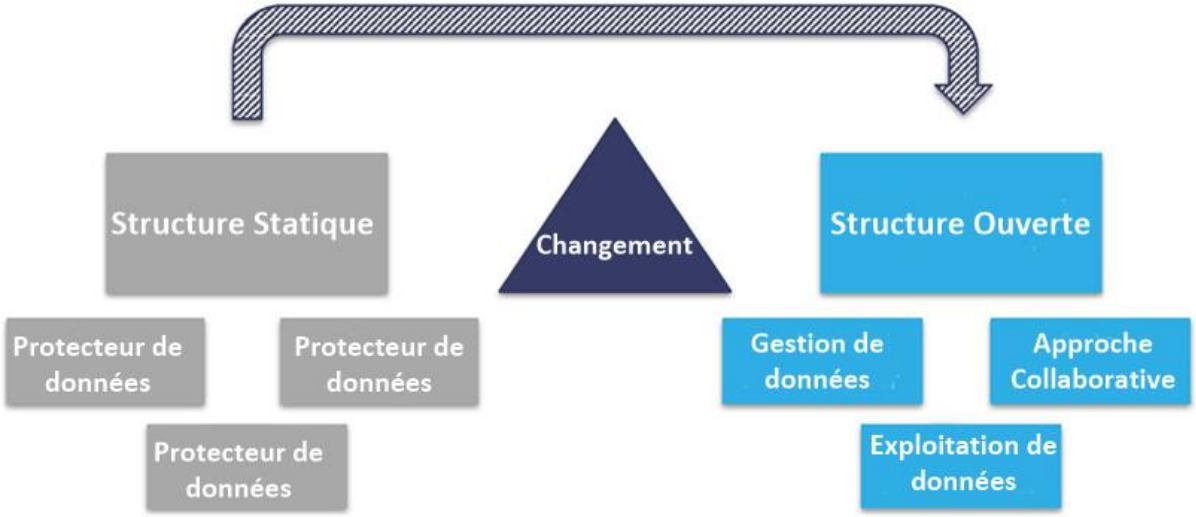


- **Des nouvelles capacités** : La mise en place de processus robustes de gestion de données, l'exploitation, le croisement, l'analyse, la visualisation et la publication de données nécessitent des compétences spécifiques qui doivent être développées au sein de la structure.
- **Des nouvelles approches collaboratives** : Le principe fondateur de l'Open Data est le rapprochement entre les structures publiques et les acteurs non-gouvernementaux. Un des objectifs de l'Open Data et de faciliter les collaborations entre les structures publiques et ces acteurs.

La mise en place concomitante ces trois éléments, supportée par une législation incitative, sont les clés de la réussite d'une initiative Open Data et apporte un ensemble de bénéfices mesurables pour les structures publiques notamment :

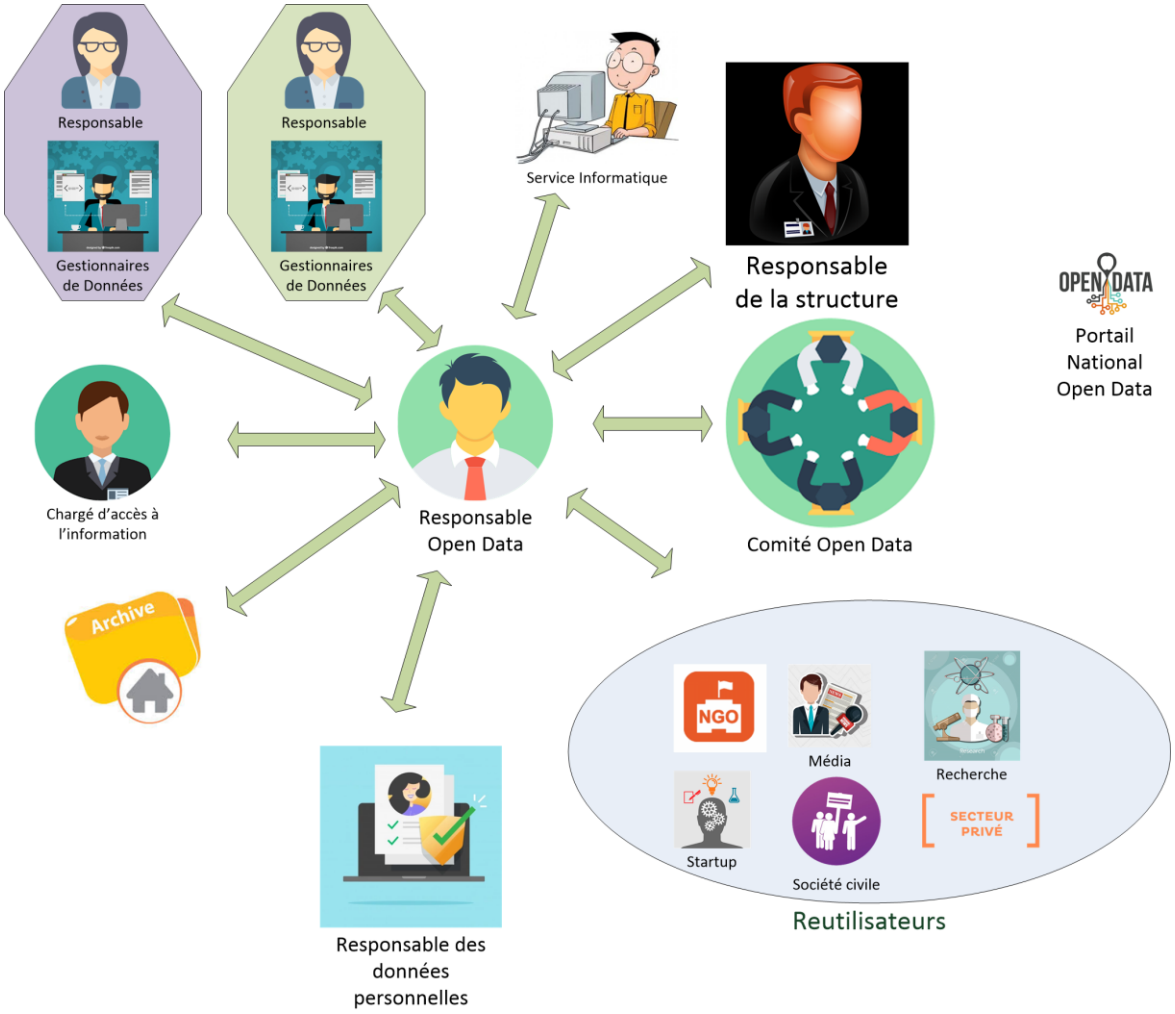
- **Economie et efficacité budgétaire** : Un des premiers impacts de la mise en place de ces approches au sein des structures concerne les économies potentielles. Par exemple, la centralisation et le partage vont faire clairement apparaitre les duplications de collecte de données qui pourraient exister au sein d'une même structure ou entre les structures.
- **Efficacité** : L'utilité principale des approches basées sur les données concerne l'amélioration de l'efficacité de l'action publique. D'une part, les politiques publiques bénéficient de prise de décision scientifique basée sur les données réelles de terrain. Ensuite, les collaborations entre les acteurs gouvernementaux et non-gouvernementaux permettent non seulement de construire un consensus partagé sur le diagnostic d'un problème donné, mais permettent également une action collective concertée sur les grands défis qui touchent le pays que ce soit lors de crises graves comme celle du Covid-19 ou sur des sujets à long terme comme la modernisation et le développement global du pays. Enfin, l'échange facilité et automatique de données entre les structures publiques peut, par exemple, identifier et éviter les dépenses inutiles, et permettre d'améliorer l'efficacité des processus et la prestation des services publics.
- **Innovation** : Une des idées fondatrices de l'Open Data est liée à la collaboration entre les structures publiques et les acteurs non-gouvernementaux qui permet l'exploitation de l'innovation citoyenne dans la résolution des problèmes de l'administration. Les approches collaboratives Open Data favorisent naturellement l'émergence de solutions innovantes issues des échanges et de la cocréation entre des acteurs ayant des points de vue complémentaires.
- **Amélioration des services publics** : Un des effets les plus visibles de la publication de données, notamment par les acteurs de l'innovation (startup…) est l'émergence de nouveaux services pour les citoyens qui viennent compléter les services publics de l'administration. D'autre part, les approches collaboratives notamment et l'organisation de compétitions par exemple permettent à l'Administration de facilement codévelopper avec ces acteurs de l'innovation de nouvelles applications innovantes et à faible coût comparé aux appels d'offre classiques.

En résumé, l'émergence d'une initiative Open Data qui produit des résultats économiques et sociaux probant requiert la transformation interne des structures publiques comme présenté dans le diagramme ci-dessous.


Structure Statique
Changement
Structure Ouverte
Protecteur de données
Protecteur de données
Protecteur de données
Gestion de données
Approche Collaborative
Exploitation de données

3. Rôles, fonctions et activités du ROD :

Dans le cadre de la conduite du changement présentée à la section précédente, le rôle du ROD est essentiel parce qu'il est à l'intersection de l'ensemble des acteurs, comme présenté dans le diagramme ci-dessous :



3.1. Les interlocuteurs du ROD :

Sous la responsabilité du responsable de la structure qui fixe la vision, la stratégie et les objectifs prioritaires, le ROD travaille en étroite collaboration avec tous les acteurs de l'Open Data impliqués dans la publication et la réutilisation de données :

- **L'Agence de Développement du Digital (ADD) et le Comité de Pilotage Open Data (CPOD)** : Le rôle de l'ADD et du CPOD est de définir, en coordination avec les acteurs concernés, le cadre réglementaire et technique de l'initiative Open Data, et de faire le suivi de l'évolution de l'implémentation au niveau de l'ensemble des organismes et institutions publics concernés où le ROD est le relais principal de l'ADD et du CPOD. D'une part, il transmet et applique les choix techniques validés par le CPOD. D'autre part, il sert de relais pour faire remonter à la fois les défis qui se posent à sa structure

et les succès rencontrés qui peuvent servir d'exemple aux autres structures. L'ADD et le CPOD sont également au support des ROD pour les aider à accomplir leurs tâches.

- **Les Gestionnaires/producteurs de données** (GdD) : Les gestionnaires de données sont les partenaires quotidiens du ROD. Il s'assure qu'ils acquièrent et exploitent les compétences nécessaires à la préparation, la publication et l'utilisation (croisement, analyse, visualisation…) des données en suivant les standards retenus à l'échelle nationale. Il contrôle également que les jeux de données publiés sont documentés et maintenus (mises à jour, réponses aux questions et commentaires, etc.).
- **Le service informatique** (SI) : Le ROD travaille en collaboration avec les GdD et le service informatique afin d'aider à l'automatisation des tâches de publication et l'intégration de la publication Open Data dans le cycle global de production des données. Comme pour les GdD, le ROD s'assure que le SI acquiert les compétences nécessaires à ces activités et utilise les standards techniques en vigueur à l'échelle nationale.
- **Le chargé d'accès à l'information** (CAI) : Le CAI et le ROD ont des missions complémentaires et des synergies existent entre ces deux fonctions. Le CAI est principalement focalisé sur le contenu, à savoir les données et documents qui peuvent être publiés ou divulgués. Le ROD est lui focalisé sur l'accessibilité des données et leur publication sous forme de données ouvertes. De ce fait, le CAI et le ROD collaborent naturellement sur des outils communs comme l'inventaire des données. L'Open Data constitue une opportunité pour le CAI et peut lui permettre de diminuer sa charge de travail notamment au travers de l'augmentation de la publication proactive qui amène mécaniquement une diminution des requêtes. De la même manière, la publication par le CAI des réponses aux requêtes sous formats ouverts favorise la réutilisation et la diminution des requêtes. Dans ce cadre, le rôle principal du ROD est de sensibiliser le CAI à ces opportunités et de lui fournir les outils et formations appropriés.
- **L'entité en charge des archives** : L'entité en charge des archives est un interlocuteur potentiel du ROD. En fonction des requêtes de demande d'accès aux archives, l'entité en charge des archives collaborera avec le ROD pour aider à la publication de jeux de données spécifiques ou pour informer les gestionnaires de données sur des droits de publications de certains jeux.
- **Le Responsable des Données Personnelles (RDP)** : Le RDP est un interlocuteur potentiel du ROD qui lui permet, notamment dans le cadre du développement d'un inventaire de données, d'identifier les jeux de données de la structure qui contiennent des données personnelles et qui doivent donc être anonymisés avant publication.
- **Les réutilisateurs** : Comme présenté à la section 1.4, il existe plusieurs catégories de réutilisateurs (société civile, média, secteur privé, acteurs de l'innovation, autres structures publiques, etc.). Le rôle du ROD est d'établir des liens et collaborer avec ces réutilisateurs.

Le ROD doit notamment :

- Identifier les réutilisateurs susceptibles d'être intéressés par les données de sa structure et parmi ces réutilisateurs identifier les leaders et les pionniers qui ont une vision constructive et sont prêts à collaborer avec l'administration
- Identifier les missions et les centres d'intérêt de ces réutilisateurs.
- Identifier les canaux de communication appropriés et les lieux d'échange de ces réutilisateurs afin d'adapter sa communication et sa stratégie d'engagement.
- Identifier les éléments qui pourraient créer une dynamique : ouverture de jeux de données spécifiques, participation à leurs événements, organisation d'événements communs, etc.

3.2. Les rôles du ROD :

Le rôle principal du ROD est de mettre en œuvre la conduite du changement de sa structure telle que décrite dans la section 2 de ce document. En particulier, il s'agit de :

- **Créer une dynamique et une émulation au sein de la structure** pour amener les gestionnaires de données à
  - Préparer et publier des données ouvertes sur le portail national Open data ou sur le portail de la structure concernée.
  - Exploiter les données de la structure et celles des autres structures pour la résolution de problèmes et la prise de décision.
- **Fournir un support aux gestionnaires de données pour faciliter leurs tâches** : Plusieurs défis limitent souvent la capacité des gestionnaires de données à accomplir leurs tâches, notamment :
  - **Les problèmes administratifs** : Une des barrières principales à la publication de données est l'aspect administratif et notamment le droit pour les gestionnaires de données de publier des jeux de données spécifiques. Un des rôles clés du ROD est de gérer et de résoudre ces problèmes administratifs afin que les GdD ne s'occupent que des aspects techniques de la publication. Une des clés dans la résolution des problèmes administratifs réside dans la mise en place d'un inventaire de données (voir la section 3.3.2.).
  - **Les standards techniques** : La deuxième difficulté des GdD est la connaissance et l'application des standards techniques de l'Open Data (format de fichier, utilisation de référentiels…). Un des rôles du ROD est d'informer et de communiquer aux GdD les standards techniques adoptés au niveau national.
  - **Les capacités techniques** : Enfin, la dernière difficulté majeure pour certains GdD réside dans l'absence de capacités techniques, soit pour mettre en œuvre les étapes de publication de données ouvertes, soit pour exploiter les données existantes. Le rôle du ROD est d'identifier l'expertise et les capacités qui font défaut aux GdD et de mettre en place les formations correspondantes.

- **L'accès aux données des autres structures** : Le ROD est le point de contact entre les structures pour l'échange de données. Il assiste les GdD qui auraient besoin de données spécifiques issues d'autres structures.

- **Créer une dynamique avec les réutilisateurs** : Les réutilisateurs ont un rôle majeur dans l'écosystème Open Data. Ils permettent, d'une part, au ROD d'identifier la demande et prioriser ses actions, et d'autre part, de créer des synergies avec l'administration et notamment de fournir un appui (innovation, compétences techniques). Un des rôles majeurs du ROD est donc d'être le point de contact et l'organisateur de cette collaboration entre la structure et les acteurs non-gouvernementaux qui sont dans son périmètre.
- **Faire le suivi, l'évaluation et le contrôle des activités** : En dehors de la promotion de l'Open Data et de la coordination des différentes activités, un des rôles majeurs du ROD consiste en une mission de contrôle notamment de la qualité des publications de jeux de données pour s'assurer que les standards techniques et administratifs (métadonnées) adoptés au niveau national, sont correctement et systématiquement appliqués. Il a également pour mission de faire un retour d'expérience aux entités chargées de l'Open Data au niveau national (CPOD, ADD), des pratiques réussies ainsi que les défis rencontrés au sein de son organisme ou institution.

La réalisation de ces objectifs s'appuie sur un ensemble d'outils qui sont présentés dans la section 3.3. Le succès de l'action du ROD se mesure principalement sur la publication et la maintenance de jeux de données sur le portail national ou le portail de la structure concernée, et l'émergence d'exemples de cas d'usage qui mettent en lumière le rôle de l'utilisation des données dans la résolution de problèmes, et les collaborations avec les réutilisateurs.

3.3. Les outils du ROD :

Afin de réaliser les objectifs présentés à la section précédente, le ROD devra disposer de plusieurs outils dont les principaux sont présentés ci-dessous.

3.3.1. La communication, le partage d'expérience, et l'émulation :

Le premier outil du ROD est la communication. La sensibilisation du personnel de la structure aux concepts de l'Open Data et à ses bénéfices pour la structure et pour le pays constitue le premier pas pour engager le changement. La dynamique au sein d'une structure s'entretient ensuite au travers de l'émulation et le partage d'expérience. Au travers de rencontres, et de communications en ligne (forums, newsletters, blogs), le ROD fera la promotion des activités en cours et favorisera l'entre-aide entre les équipes. Cette approche devra être déclinée en interne pour motiver les équipes et en public en tant que moyen d'interaction et d'engagement des réutilisateurs.

3.3.2. L'inventaire de données :

L'inventaire de données est un élément important de l'écosystème Open Data et un guide essentiel pour les gestionnaires de données. D'une part, il permet de connaître l'ensemble des jeux de données disponibles au sein d'une structure, ainsi que leurs niveaux de confidentialité et les traitements nécessaires à appliquer avant leur publication. D'autre part,

il permet également de détecter les duplications d'efforts dans la collecte de données entre les structures, et facilite donc la réutilisation entre structures. Enfin, c'est un outil qui résout l'ensemble des questions administratives qui entravent la publication, et qui permet aux GdD de se focaliser uniquement sur les aspects techniques de la publication. L'inventaire répond à toutes les questions notamment de droit d'accès qui facilitent la résolution des questions relatives à l'accès public.

L'inventaire de données est également un outil important pour le ROD et pour la collaboration avec les réutilisateurs. En effet, l'inventaire identifie de façon claire les jeux de données qui sont publiés, les jeux de données qui peuvent être publiés et les jeux de données qui ne peuvent pas être rendus publiques. A partir de ces éléments, les réutilisateurs et le ROD peuvent identifier les priorités concernant les jeux qui peuvent être publiés.

Il est important de noter que la réalisation d'un inventaire est une tâche de grande envergure qui doit être implémentée en plusieurs étapes et nécessite une expertise spécifique. Les moyens, approches et méthode de développement d'un inventaire sont en dehors du cadre de ce manuel. L'Annexe A de ce document présente en détail l'approche. Le rôle du ROD concernant l'inventaire est d'en coordonner le développement et la maintenance.

3.3.3. Les formations :

Le troisième outil du ROD, ce sont les formations. L'organisation de formations est essentielle pour apporter les capacités nécessaires aux différents personnels. On identifie en général deux familles de formations avec chacune des modules spécifiques :

- **Les Formations non-techniques :** Les formations non-techniques incluent au minimum les modules suivants :
    - **Formation générale à l'Open Data :** Ces modules s'adressent à tous les personnels et leurs présentent les concepts clés de l'Open Data ainsi que les bénéfices et les opportunités.
    - **Aspects légaux de l'Open Data** : Il s'agit de former les personnels aux différentes réglementations en relation avec la publication de données. Ces modules s'adressent à tous les personnels.
    - **Conduite du changement** : Ces modules s'adressent principalement aux responsables (chef de départements, etc.) et leur donnent les outils pour mobiliser leurs personnels et les approches à mettre en œuvre.
- **Formations techniques** : Les formations techniques incluent au minimum les modules suivants :
    - **Préparation et Publication de données** : Ces modules permettent aux GdD de maitriser toutes les étapes de la préparation des données jusqu'au téléchargement sur le portail Open Data.
    - **Gestion de données** : Ces modules permettent aux GdD d'acquérir les compétences nécessaires à l'exploitation des données (analyse, croisement, visualisation, etc.).

- **Gestion de portail** : Ces modules s'adressent principalement aux administrateurs systèmes des services informatiques et ont pour but de leur permettre de maitriser l'installation, la configuration et la gestion d'un portail de données ouvertes, dans le cas où la structure désire mettre en œuvre son propre portail.
- **Inventaire de données** : Ces modules s'adressent principalement aux ROD et aux services informatiques et ont pour but de leur enseigner les bases de la mise en place d'un inventaire de données.

Le rôle principal du ROD est de mobiliser les participants avec le profil adéquat et la motivation nécessaire.

3.3.4. Les cas d'usage :

Le développement de capacités et les formations théoriques sont indispensables mais non suffisantes pour que les GdD et les décideurs changent leurs mécanismes décisionnels et adoptent des démarches basées sur les données. Ce changement d'approche nécessite le plus souvent une démonstration pratique de l'intérêt de ces approches. Pour ce faire, une approche intéressante consiste à développer un exemple visible de cas d'usage. L'idée est de développer et de documenter un exemple de bout en bout depuis l'identification d'une problématique jusqu'à sa résolution au travers d'une approche basée sur les données. Le développement de ce type de cas d'usage suit généralement un processus par étape résumé dans le diagramme ci-dessous.

Ces approches permettent à la fois de mettre en œuvre les compétences techniques acquises lors des formations et faire prendre conscience aux personnels qui participent au développement de l'intérêt et de la puissance de ces approches. Elles permettent également de développer du matériel de communication qui sert à motiver les personnels et départements non encore impliqués. La réalisation de ces cas d'usage est généralement plus

fructueuse si elle est conduite en équipe avec des compétences diverses (gestionnaires de données, statisticiens, décideurs, informaticien). Le Rôle du ROD dans le développement de cas d'usage est d'organiser les équipes, d'identifier les participants et de coordonner la réalisation.

3.3.5. Les événements d'innovation, hackathons, et compétitions :

L'émulation et la création d'une dynamique peuvent être supportées par la mise en place de compétitions et d'événements d'innovation. Ces compétitions peuvent être de différents types et amènent différents bénéfices :

- **Les compétitions internes** impliquent uniquement des participants d'une même structure (ou de différentes structures mais uniquement gouvernementaux). Ces compétitions sont généralement motivantes et génératrices de résultats (cas d'usage) rapidement. Elles permettent aux participants ainsi qu'aux responsables de réaliser les bénéfices des approches Open Data. Elles permettent également d'identifier les personnels les plus motivés. Plusieurs types de compétitions peuvent être imaginées comme la réalisation de cas d'usage ou des compétitions de publication de données.
- **Les compétitions externes** (hackathon) impliquent uniquement des réutilisateurs. Elles permettent pour les structures de réaliser l'impact des publications de données et de découvrir les services innovants qui peuvent être dérivés. Ce type de compétitions génère le plus souvent une forte motivation au sein des structures lorsqu'elles réalisent le potentiel de leurs données pour le pays dans son ensemble. Ces compétitions permettent également la collaboration entre la structure et les réutilisateurs.
- **Les compétitions mixtes** (interne/externe) impliquent à la fois des personnels des organismes et institutions publics et des acteurs non-gouvernementaux. L'idée de ces compétitions est de constituer des équipes mixtes pour une réalisation conjointe de cas d'usage. Ces compétitions permettent à l'administration d'utiliser l'innovation et les compétences citoyennes pour la résolution de leurs problèmes. Ce type de compétitions renforce les liens entre l'administration et les acteurs non-gouvernementaux et permet de mettre en place des collaborations potentiellement pérennes. Ce type de compétitions est une approche innovante pour l'externalisation de certaines tâches (développement d'applications, conseil pour la résolution de problèmes spécifiques, etc.).

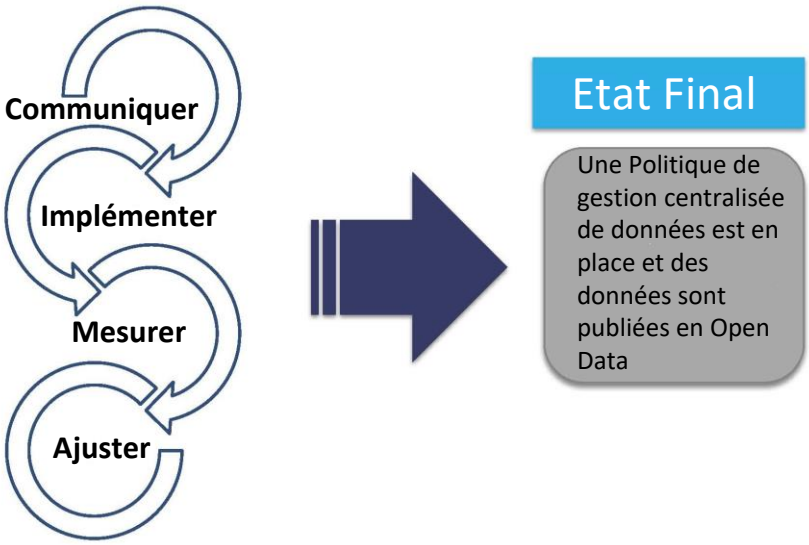
Les outils présentés ci-dessus sont des éléments unitaires. Le défi et la tâche principale du ROD est d'orchestrer ces activités afin d'accompagner le changement de la structure et atteindre les objectifs visés. La section 3.4 présente les approches à mettre en œuvre pour ce faire.

3.4. Implémenter une Approche Open Data :

Quand il s'agit de conduite du changement, dans toutes les structures, il est souvent possible de répartir les personnels en trois catégories :

- **Les innovateurs** qui sont attirés et motivés par la mise en place d'une nouvelle approche et de nouvelles façons de travailler.
- **La majorité suiveuse** qui, une fois qu'elle a vu les avantages pratiques d'une nouvelle approche, est prête à la mettre en œuvre.
- **Les réfractaires** qui sont plutôt fermés à l'idée de changer de méthodes de travail, de changer de poste ou d'acquérir de nouvelles capacités.

On retrouve ces trois catégories quand il s'agit de mettre en place une démarche Open Data. Le processus de changement doit donc s'appuyer sur cette stratification pour créer une démarche qui va intégrer ces différents groupes au fur et à mesure de l'évolution de l'initiative pour aboutir à la transformation globale de la structure. L'approche se fait donc en spirale en se focalisant d'abord sur les innovateurs qui vont démontrer la valeur et les intérêts de l'Open Data et entrainer graduellement l'ensemble de leurs collègues. Cette approche est résumée dans le diagramme ci-dessous.



Le ROD va donc développer cette approche à la fois auprès des gestionnaires de données mais également avec les réutilisateurs. Cette approche utilise deux éléments spécifiques :

- **Le plan annuel d'activité** : A partir des priorités et des objectifs fixés par le responsable de la structure, et du budget alloué, le ROD identifie les actions à mener et leur ordonnancement au cours de l'année.
- **Le plan de suivi et évaluation** : Le plan de suivi et évaluation (S&E) permet de mesurer l'impact des différentes activités, le développement de l'Open Data au sein de la structure, et les défis qu'il convient de résoudre.

3.4.1. Plan annuel d'activité Open Data :

Le plan annuel d'activité est la feuille de route de la structure concernant l'implémentation de l'ouverture des données. Ce plan est développé par le ROD en collaboration avec tous les

acteurs de l'Open Data (voir section 3.1) et validé par le responsable de la structure. Un plan annuel d'activité contient généralement les sections suivantes :

- **Objectifs** : Un des éléments clés du plan annuel d'activité est de fixer un ensemble d'objectifs spécifiques. Ces objectifs sont en général définis par le responsable de la structure. Ils peuvent être de différentes natures et complémentaires, par exemple :
    - **Thématiques** : Focalisation du plan sur un sujet donné, ou une problématique de la structure.
    - **Organisationnels** : Focalisation sur un département, ou sur une audience particulière (certains types de réutilisateurs, etc.).
    - **Fonctionnels :** Focalisation du plan sur la mise en place d'éléments spécifiques comme l'inventaire de données ou un portail de données dédiés à la structure.
    - **Quantitatifs** : Focalisation du plan sur la publication d'un certain nombre de jeux de données ou de cas d'usage.
- **Communication** : Le plan doit inclure des actions de communication interne (vers les membres de ls structure) et externe (réutilisateurs d'autres structures ou non-gouvernementaux). En particulier, il est important que la présentation du plan d'activité lui-même fasse l'objet de communication interne et externe, de préférence par le responsable de la structure, afin à la fois de mobiliser les personnels et les réutilisateurs potentiels. Enfin, la section communication peut éventuellement inclure des aspects réglementaires et administratifs (notes de service, directives, nominations officielles…).
- **Formations** : Le plan d'activité inclut à la fois les formations qui seront organisées et les cibles de ces formations en fonction des objectifs visés.
- **Inventaire de données** : La réalisation de l'inventaire de données étant une tâche exécutée par étape, le plan d'activité peut inclure l'extension de l'inventaire (ou sa mise en place).
- **Publication de données** : Le plan d'activité inclut généralement un engagement quantitatif et qualitatif sur la publication de nouveaux jeux de données.
- **Réutilisation** : Le plan contient une section qui décrira la stratégie pour stimuler la réutilisation à la fois en interne (développement de cas d'usage) et en externe (actions pour collaborer avec les réutilisateurs externes à la structure, événements, etc.)
- **Infrastructure** : Le plan d'activité peut inclure la mise en place d'infrastructures spécifiques comme un portail de données ouvertes spécifique à la structure qui sera connecté au portail national.
- **Budget** : Certaines des activités ci-dessus nécessitent des budgets spécifiques (par exemple l'organisation de compétitions, de formations, etc.). Le ROD devra établir un budget en fonction des activités sélectionnées dans le plan. A noter que certaines activités (formations, compétitions externes ou mixtes…) seront probablement organisées par l'ADD et le Comité de Pilotage Open Data et leurs coûts ne seront pas supportés par les structures directement.

Le contenu du plan annuel d'activité Open Data doit être influencé par trois éléments :

- **La stratégie globale de la structure et ses objectifs propres** (indépendamment des objectifs Open Data).
- **La demande des réutilisateurs.**
- **L'état du développement de l'initiative Open Data au sein de la structure**. Il s'agit d'identifier les défis et mesurer l'évolution de l'initiative en fonction des objectifs du plan précédent et des résultats obtenus. Pour ce faire, une des tâches du ROD consiste à maintenir un plan de suivi et évaluation présenté dans le chapitre 3.4.2.

3.4.2. Plan de Suivi et Evaluation :

Le rôle du plan de suivi et évaluation Open Data de la structure a pour objectif de mesurer l'état d'avancement de la structure dans sa conduite du changement. Le plan de suivi et évaluation contient en général les éléments suivants :

- **Des indicateurs quantitatifs automatiques** : un ensemble d'indicateurs peuvent être automatiquement renseignés à partir notamment du portail national de données (par exemple le nombre de jeux de données publiés, leur qualité en termes de mises à jour, métadonnées, format, etc., ou le nombre de téléchargement).
- **Des indicateurs liés aux activités** : Le plan de S&E doit maintenir à jour les indicateurs liés aux activités comme par exemple le nombre de participants aux formations, le nombre de cas d'usage publiés, ou le nombre d'ateliers organisés.
- **Des indicateurs de suivi** : Le plan de S&E contient un ensemble d'indicateurs qui reflète la qualité de maintenance des jeux de données publiés, notamment la régularité de leur mise à jour, la réponse aux commentaires des réutilisateurs, le taux de réutilisation des données publiées, et le nombre d'applications/ services crées en utilisant les données ouvertes publiées.

Ces indicateurs sont généralement répartis en plusieurs sections qui reflètent le contenu du plan d'activité. On trouve en général les sections suivantes :

- **Gouvernance** : cette section regroupe les indicateurs liés à la gouvernance (par exemple mise en place d'un plan annuel d'activité, d'un plan de suivi et évaluation, etc.)
- **Qualité des jeux de données publiées** : Cette section regroupe les indicateurs qui reflètent de la qualité des jeux de données publiés.
- **Publication** : Cette section regroupe les indicateurs qui reflètent du volume de jeux de données publiés.
- **Réutilisation** : Cette section regroupe les indicateurs qui reflètent le dynamisme et le volume de réutilisation interne et externe des données publiées.

- **Portail de données** : Cette section, présente uniquement si la structure possède son propre portail, regroupe les indicateurs liés au portail.

Normalement, le contenu du plan de suivi et évaluation (indicateurs) est standardisé au sein de l'administration afin de faciliter le suivi global par le Comité de Pilotage Open Data. Un exemple de plan de suivi et évaluation est fourni en annexe B.

3.5 Les écueils à éviter :

Cette section rassemble les écueils à éviter pour les ROD dans la réalisation de leur tâche. Cette section sera enrichie au fur et à mesure des retours d'expériences et des évaluations qui seront menées.

- **Le volume de publication (nombre de jeux de données) n'est pas le facteur le plus important** : Le volume de publication d'une structure est rarement un indicateur pertinent pour mesurer l'activité Open Data. Il est en effet relativement facile pour une structure de publier un grand nombre de jeux de données qui ont peu de valeur parce qu'ils intéressent peu ou pas les réutilisateurs. Il est essentiel de focaliser la publication de données sur :
    - Des données à forte valeur, à savoir des données en demande des réutilisateurs.
    - Des données de qualité publiées dans des formats exploitables et documentés.
    - Des données maintenues, qui sont régulièrement mises à jour en fonction du cycle de collecte, et sur lesquels les GdD apportent un support (réponses aux commentaires et aux questions, corrections).

  Il est à souligner également l'importance de la documentation des données à publier notamment à travers la description et les mots clés qui permettent une meilleure interprétation et analyse de ces données par les utilisateurs.

- **Une fois que l'ensemble des jeux de données seront publiés, le travail est loin d'être terminé** : La publication de données n'est que la première étape et le premier étage du processus Open Data. L'objectif principal est d'une part de mettre la donnée au cœur de l'action publique et des structures, et d'autre part de valoriser les données de l'administration afin de générer un impact social et économique. Si la publication est un élément essentiel, il est tout aussi important de se focaliser sur la maintenance des jeux de données, et sur la réutilisation de ces données en interne ou en externe. Enfin, il est important de noter que l'action du ROD n'est pas une action limitée dans le temps. La gestion de données au sein d'une structure est un processus dynamique où la collecte de certaines données s'arrête et où des nouveaux jeux de données apparaissent. Le processus de contrôle et de suivi du patrimoine de données d'une structure est donc une tâche qui est pérenne dans le temps. Comme on peut l'observer dans les pays les plus avancés sur le sujet, à terme, le rôle du ROD évoluera naturellement vers une fonction dédiée à la gestion de l'inventaire et son suivi

(administrateur général des données ou « Chief Data Officer[11] » dans les pays anglo-saxon).

[11] https://fr.wikipedia.org/wiki/Chief_Data_Officer

Annexe A – Inventaire de données :

L'inventaire de données est un élément important de l'écosystème Open Data et un guide essentiel pour les producteurs de données. D'une part il permet de connaître l'ensemble des jeux de données disponibles au sein d'une structure, ainsi que leurs niveaux de confidentialité et les traitements nécessaires à appliquer avant leur publication. D'autre part, il permet également de détecter les duplications d'efforts dans la collecte de données entre les structures, et facilite donc la réutilisation entre structures. Enfin, c'est un outil qui résout l'ensemble des questions administratives qui entravent la publication.

La réalisation d'un inventaire complet est une tâche de grande envergure qui devrait être implémentée en plusieurs étapes. Dans un premier temps, une version initiale serait développée en partant de l'existant et en incluant tous les ministères précurseurs. Pour cette première version, une méthodologie serait établie. Ensuite, chaque nouvelle structure rejoignant l'initiative d'ouverture des données appliquerait la méthodologie définie. Afin de démarrer le processus dans toutes les structures et ministères, il serait important, au travers du décret Open Data, de forcer toutes les entités à avoir un inventaire minimum contenant les données dont la publication est rendue obligatoire par la loi d'accès à l'information.

Concernant la méthodologie, il faudra dans un premier temps définir un schéma directeur des données à collecter (structure en charge, personne en charge, périodicité de la collecte, stockage, format, description, traitement à appliquer avant publication, niveau de confidentialité, lien vers les données publiées, niveau de maturité des jeux de données par rapport à la classification défini par Sir Tim Berners-Lee, inventeur du Web[12], etc.). Une fois ce schéma établi, les informations pourront être collectées.

Il conviendra d'adopter une démarche structurée pour identifier les jeux de données. Cette démarche devra au minimum inclure l'exploration des éléments suivants :

- Les sites Web de la structure.
- Les systèmes d'information, s'il y en a.
- Les observatoires associés à la structure s'il y en a : les observatoires ont pour mission de collecter un grand nombre de données pour construire leurs indicateurs.
- Les données fournies au HCP.

A noter qu'il est presque impossible d'être exhaustif dans la liste des données collectées par une structure. La version initiale permet de fournir un premier aperçu. Ensuite, cette information sera complétée en fonction des demandes reçues de l'extérieur via les chargés d'accès à l'information.

Enfin, il est essentiel de noter qu'un inventaire est une information vivante qui évolue dans le temps. Certains jeux arrêtent d'être collectés, certains autres sont ajoutés. Si la réalisation de la version initiale de l'inventaire peut être externalisée, il sera important pour la pérennité

[12] http://5stardata.info/fr/

de cette information que, dans chaque ministère, une personne soit en charge de sa gestion et de sa maintenance. Cette personne pourra être soit le Directeur des Systèmes d'Information s'il existe, soit le chargé d'accès à l'information, soit le chargé de l'Open Data.

Références :

Les ressources suivantes fournissent les grandes lignes du développement d'un inventaire et des exemples d'inventaires de différents pays.

- Guide de l'inventaire des Etats-Unis[13].
- Guide de l'inventaire de la Confédération Helvétique[14].
- Inventaire du Canada[15].
- Guide de l'Inventaire de la ville de Washington DC[16].
- Guide pour la réalisation d'un inventaire de données du Centre pour l'excellence dans la gouvernance (GovEx lab) de l'université Johns Hopkins (anglaise)[17].
- Guide pour la réalisation d'un inventaire de données – Initiative Open Data de la ville de San Francisco (anglais)[18].

[13] https://project-open-data.cio.gov/implementation-guide/
[14] https://handbook.opendata.swiss/en/identify/inventory.html
[15] https://open.canada.ca/data/en/dataset/4ed351cf-95d8-4c10-97ac-6b3511f359b7
[16] https://opendata.dc.gov/datasets/76a28737a6f84b3c92a421114acccca2_5
[17] http://labs.centerforgov.org/data-governance/data-inventory/
[18] https://datasf.org/resources/data-inventory-guidance/

Annexe B – Exemple de plan de suivi et évaluation :

Catégorie	Indicateur	Type	Description	Mesure	Valeur minimale	Source
Gouvernance	Publication du plan de suivi & évaluation de l'organisme	URL	URL vers la grille d'évaluation	Le jeu de données correspondant sur le portail est à jour	Oui (obligation)	Portail national de données
	Publication du rapport de synthèse de suivi de l'initiative et des organismes	URL	URL vers le rapport	Une URL pointant sur le rapport de l'année existe	Oui (obligation)	Site Web de l'organisme
	Publication du plan annuel d'activité	URL	URL vers le plan		Oui (obligation)	Site Web de l'organisme
	% des activités prévues dans le plan d'action réalisées	%	Rapport entre le nombre d'activités implémentées sur l'année et le nombre total d'activités contenues dans le plan	Mesure manuelle basée sur l'analyse du rapport d'activité vs le plan annuel d'activité	Objectif 100%	Rapport annuel d'activité - Plan annuel d'activité
	Budget alloué aux activités Open Data	Budget		Information fournie manuellement sous la forme d'une liste de lignes budgétaires finançant le différent type d'activités Open Data		Rapport annuel d'activité

	% des jeux de données inventoriés dans l'inventaire ayant l'ensemble des métadonnées requises pour l'inventaire	%	Critère de qualité de l'inventaire et de respect du modèle adopté par le CPOD	Tous les inventaires sont publiés sous format ouvert et peuvent être analysés		Portail de l'organisme (à défaut, portail national)
Qualité des jeux de données	% des jeux ayant des métadonnées	%		Analyse automatique des jeux de données publiées	100% obligation de publication	Portail de l'organisme (à défaut, portail national)
	% des jeux ayant les métadonnées recommandées	%		Analyse automatique des jeux de données publiées	100% obligation de publication	Portail de l'organisme (à défaut, portail national)
	% des jeux de données ayant une documentation	%		Analyse automatique des jeux de données publiées	100% obligation de publication	Portail de l'organisme (à défaut, portail national)
	% des jeux de données utilisant un format recommandé par le Comité de Pilotage Open Data	%		Analyse automatique des jeux de données publiées	100% obligation de publication	Portail de l'organisme (à défaut, portail national)
	% des jeux de données publié sous plusieurs formats	%		Analyse automatique des jeux de données publiées	Pas de valeur minimale	Portail de l'organisme (à défaut, portail national)

	% des jeux de données atteignant le niveau 3* de l'échelle de Tim Berners-Lee	%		Analyse automatique des jeux de données publiées	100% pour les jeux de données publiés après l'entrée en vigueur du décret	Portail de l'organisme (à défaut, portail national)
	% des jeux de données ayant une licence	%		Analyse automatique des jeux de données publiées	100% obligation de publication	Portail de l'organisme (à défaut, portail national)
	% des jeux de données utilisant la licence officielle adoptée par le CPOD ou une homologuée	%		Analyse automatique des jeux de données publiées	100% obligation de publication	Portail de l'organisme (à défaut, portail national)
	% des jeux de données utilisant une licence non homologuée par le CPOD	%		Analyse automatique des jeux de données publiées	0% (valeur maximale)	Portail de l'organisme (à défaut, portail national)
	% des jeux de données publiés automatiquement sur le portail	%		Analyse automatique des jeux de données publiées	Pas de valeur minimale	Inventaire de données de l'organisme
	% des jeux de données utilisant un ou plusieurs référentiels officiels adoptés par le CPOD	%		Analyse automatique des jeux de données publiées	Pas de valeur minimale	Portail de l'organisme (à défaut, portail national)

	% des jeux de données utilisant un référentiel de données périmé	%		Analyse automatique des jeux de données publiées	0% (valeur maximale)	Portail de l'organisme (à défaut, portail national)
	% des jeux de données présentant des données désagrégées jusqu'au secteur	%		Analyse automatique des jeux de données publiées	Pas de valeur minimale	Portail de l'organisme (à défaut, portail national)
	Nombre de jeux de données à jour	Chiffre		Analyse automatique des jeux de données publiées	100% obligation de publication	Portail de l'organisme (à défaut, portail national)
	Nombre de jeux de données nécessitant une mise à jour	Chiffre		Analyse automatique des jeux de données publiées	0% (valeur maximale)	Portail de l'organisme (à défaut, portail national)
	Nombre de jeux de données périmés	Chiffre		Analyse automatique des jeux de données publiées	0% (valeur maximale)	Portail de l'organisme (à défaut, portail national)
Publication	Nombre de jeux de données inventoriés	Chiffre		Analyse du jeu de données de l'inventaire	Pas de valeur minimale	Inventaire de données de l'organisme
	Nombre de jeux de données classifiés comme publiables	Chiffre		Analyse du jeu de données de l'inventaire	Pas de valeur minimale	Inventaire de données de l'organisme

	Nombre de jeux de données publiés	Chiffre		Analyse du catalogue du portail national	Pas de valeur minimale	Portail de l'organisme (à défaut, portail national)
	Nombre de requêtes de publications reçues	Chiffre		Information suivie par le CAI	Pas de valeur minimale	CAI
	Nombre de requêtes de publications répondues	Chiffre		Information suivie par le CAI	Pas de valeur minimale	CAI
	Nombre de requêtes de publications satisfaites	Chiffre		Information suivie par le CAI	Pas de valeur minimale	CAI
	Nombre de requêtes de modification/mise-à-jour reçues	Chiffre		Analyse des commentaires	Pas de valeur minimale	Portail de l'organisme (à défaut, portail national)
	Nombre de requêtes de modification/mise-à-jour reçues répondues	Chiffre		Analyse des commentaires	Pas de valeur minimale	Portail de l'organisme (à défaut, portail national)
	Nombre de requêtes de modification/mise-à-jour reçues satisfaites	Chiffre		Analyse des commentaires	Pas de valeur minimale	Portail de l'organisme (à défaut, portail national)

	Publication de l'ensemble minimum des jeux de données prévu par la loi d'accès	Oui/Non		Indicateur rempli manuellement	Oui - obligation de la loi d'accès	Portail de l'organisme (à défaut, portail national)
Formation	Nombre de sessions de formation technique organisées	Chiffre		Agrégation des informations par organisme plus les sessions spécifiques organisées par le CPOD	Pas de valeur minimale	Rapport annuel d'activité
	Nombre de sessions de sensibilisation/conduite du changement organisées	Chiffre		Agrégation des informations par organisme plus les sessions spécifiques organisées par le CPOD	Pas de valeur minimale	Rapport annuel d'activité
	% des hauts responsables formés	%		Agrégation des informations des organismes	Pas de valeur minimale	Rapport annuel d'activité
	% des gestionnaires de données formés	%		Agrégation des informations des organismes	Pas de valeur minimale	Rapport annuel d'activité
	% des CAI formés	%		Agrégation des informations des organismes	Pas de valeur minimale	Rapport annuel d'activité
	% des ROD formés	%		Agrégation des informations des organismes	Pas de valeur minimale	Rapport annuel d'activité
	% des informaticiens formés	%		Agrégation des informations des organismes	Pas de valeur minimale	Rapport annuel d'activité
Réutilisation	Nombre de jeux de données uniques téléchargés par des membres de l'administration	Chiffre		Exploitation des logs du portail (national ou spécifique)	Pas de valeur minimale	Web analytics du portail de données de l'organisme ou national

	Nombre de jeux de données uniques téléchargés par des acteurs non-gouvernementaux	Chiffre		Exploitation des logs du portail (national ou spécifique)	Pas de valeur minimale	Web analytics du portail de données de l'organisme ou national
	Nombre de téléchargements par des membres de l'administration	Chiffre		Exploitation des logs du portail (national ou spécifique)	Pas de valeur minimale	Web analytics du portail de données de l'organisme ou national
	Nombre de téléchargements par des acteurs non-gouvernementaux	Chiffre		Exploitation des logs du portail (national ou spécifique)	Pas de valeur minimale	Web analytics du portail de données de l'organisme ou national
	Nombre d'activités de promotion de la réutilisation organisées (hackathon, sensibilisation, etc.)	Chiffre		Collecte manuelle	Pas de valeur minimale	Rapport annuel d'activité
	Nombre total d'exemples de réutilisations	Chiffre		Exploitation des informations stockées sur le portail	Pas de valeur minimale	Portail de données de l'organisme ou national

	Répartition sectoriel et géographique des téléchargements	Chiffre		Exploitation des informations collectées sur les réutilisateurs (cf. la stratégie du CPOD sur le sujet)	Pas de valeur minimale	Web analytics du portail de données de l'organisme ou national et informations collectées (utilisateurs, enquêtes etc.)
Portail	Utilisation du portail national	Oui/Non		Indicateur rempli manuellement	Si oui, les indicateurs ci-dessous sont inutiles	NA
	Interconnexion avec le portail national	Oui/Non		Indicateur rempli manuellement		Portail national de données
	Utilisation de la plateforme technique recommandée par le CPOD	Oui/Non		Indicateur rempli manuellement		Portail de données de l'organisme
	Nombre de liens cassés	Chiffre		Exploitation des logs des portails	0% (persistance des url)	Web analytics du portail de données
	Nombre d'utilisateurs uniques sur le portail	Chiffre		Exploitation des informations stockées sur les portails	Pas de valeur minimale	Web analytics du portail de données
	Nombre d'utilisateurs uniques de l'api	Chiffre		Exploitation des informations stockées sur les portails	Pas de valeur minimale	Web analytics du portail de données
	Nombre d'utilisateurs enregistré	Chiffre		Exploitation des informations stockées sur les portails	Pas de valeur minimale	Web analytics du portail de données

	Nombre d'utilisateurs uniques laissant des commentaires	Chiffre		Exploitation des informations stockées sur les portails	Pas de valeur minimale	Web analytics du portail de données
	Nombre d'utilisateurs uniques publiant des exemples de réutilisations	Chiffre		Exploitation des informations stockées sur les portails	Pas de valeur minimale	Web analytics du portail de données
	Nombre de producteurs de données	Chiffre		Analyse des jeux de données	Pas de valeur minimale	Portail de données
	Nombre de téléchargements de jeux de données via l'api	Chiffre		Exploitation des logs des portails	Pas de valeur minimale	Web analytics du portail de données
	Nombre de téléchargements de jeux de données directement	Chiffre		Exploitation des logs des portails	Pas de valeur minimale	Web analytics du portail de données