

Volet Information

Sous-arbre sur les données ouvertes

Table des matières

LISTE DES FIGURES	VI
LISTE DES SIGLES ET ACRONYMES	VII
HISTORIQUE DES CHANGEMENTS	VIII
AVIS AUX LECTEURS	IX
1. INTRODUCTION	1
1.1 LES DONNÉES OUVERTES DANS LE CONTEXTE D'UNE SOLUTION D'AFFAIRES	1
2. SOUS-ARBRE SUR LES DONNÉES OUVERTES	2
2.1 SOUS-ARBRE (VOLET INTRANT)	3
2.2 DES DONNÉES PEUVENT-ELLES ÊTRE OBTENUES DE DONNÉES QUÉBEC?	3
2.3 DES DONNÉES DE NATURE PUBLIQUE POURRAIENT-ELLES ÊTRE OBTENUES D'UNE ORGANISATION?	4
2.4 SOUS-ARBRE (VOLET EXTRANT)	4
2.5 IDENTIFIER LES INFORMATIONS CANDIDATES À ÊTRE DIFFUSÉES PUBLIQUEMENT	6
2.6 RESTRICTIONS DE DIFFUSIONS?	7
2.7 PRIORISER LES DONNÉES CANDIDATES À LA DIFFUSION	8
2.8 INSCRIRE LES DONNÉES AU CALENDRIER DE DIFFUSION	8

Liste des figures

FIGURE 1 : SOUS-ARBRE SUR LES DONNÉES OUVERTES (VOLET INTRANT)	3
FIGURE 2 : SOUS-ARBRE SUR LES DONNÉES OUVERTES (VOLET EXTRANT)	5

Liste des sigles et acronymes

API	Application programming interface
DGEQ	Directeur général des élections du Québec
PRP	Protection des renseignements personnels
SCT	Secrétariat du Conseil du trésor
TI	Technologies de l'information

Historique des changements

Version	Date de publication	Modifications
4.0	Juin 2018	Publication de la première édition

Avis aux lecteurs

Le sous-arbre présenté dans ce document détaille le volet Données ouvertes de l'arbre de décision général. Le lecteur est donc invité à prendre connaissance de l'arbre de décision général afin d'obtenir un aperçu complet de la démarche.

1. Introduction

Le document Sous-arbre sur les données ouvertes accompagne le lecteur dans l'application des bonnes pratiques en matière de données ouvertes dans un projet en technologies de l'information (TI).

En se référant aux deux sous-arbres présentés dans ce document, le lecteur sera en mesure de cibler les données qu'il peut utiliser pour développer son projet en TI (données en intrant) et également celles à diffuser en vue de leur réutilisation par divers partenaires (données en extrant).

1.1 Les données ouvertes dans le contexte d'une solution d'affaires

L'information gouvernementale est un actif stratégique de l'administration publique. Son organisation cohérente et sa gestion efficace contribuent à l'amélioration de la prestation de services pour les citoyens et les entreprises.

Bien souvent, une même information est utile à plus d'un organisme public pour qu'il réalise sa mission. La pratique courante est que chaque organisme public collecte et gère cette information selon ses propres moyens. Cette façon de procéder laisse place à un risque élevé d'incohérence et de perte d'intégrité des données.

Dans ce contexte, la diffusion de l'information gouvernementale en données ouvertes permet d'en assurer une meilleure cohérence auprès des organismes publics et des partenaires externes comme les entreprises, les centres de recherche, le public, etc.

Plus concrètement, l'utilisation et la diffusion de données ouvertes comportent les avantages suivants :

- **Une licence** qui garantit la réutilisation des données sans entente particulière.
- **Des formats ouverts** qui permettent une interopérabilité entre les différents systèmes.
- **Une disponibilité** des données en temps réel, notamment par l'utilisation des API offerts par Données Québec.

Que sont les données ouvertes?

Les données ouvertes sont des données brutes, sous la forme de fichiers numériques, diffusées de manière structurée selon une méthode garantissant leur libre accès et leur réutilisation en vertu des conditions permises par leur licence.

Au Québec, les données ouvertes sont diffusées sur le portail [Données Québec](#). Plus de 30 organismes publics y diffusent plus de 300 jeux de données. En considérant la dizaine de municipalités partenaires, plus de 1000 jeux de données y sont diffusés.

À titre d'exemple, voici des données ouvertes réutilisées par les organismes publics :

- [Répertoire des municipalités](#);
- [Registre des entreprises](#);
- [Adresses Québec](#).

Voici d'autres exemples de données spécifiques à des domaines d'affaires et réutilisées par la société civile :

- [Carte écoforestière avec perturbations](#);
- [Base de données des zones inondables](#).

2. Sous-arbre sur les données ouvertes

Cette section présente les deux sous-arbres sur les données ouvertes :

- **Sous-arbre (volet intrant)**
- Ce sous-arbre permet de cibler les données ouvertes à obtenir pour qu'elles soient utilisées dans un système d'information.
- **Sous-arbre (volet extrant)**
- Ce sous-arbre permet de cibler les données à extraire d'un système d'information pour qu'elles soient diffusées en données ouvertes.

Rappel sur les étapes réalisées antérieurement dans l'arbre de décision général.

En ayant suivi l'ensemble des étapes de l'arbre de décision général, vous êtes en mesure de :

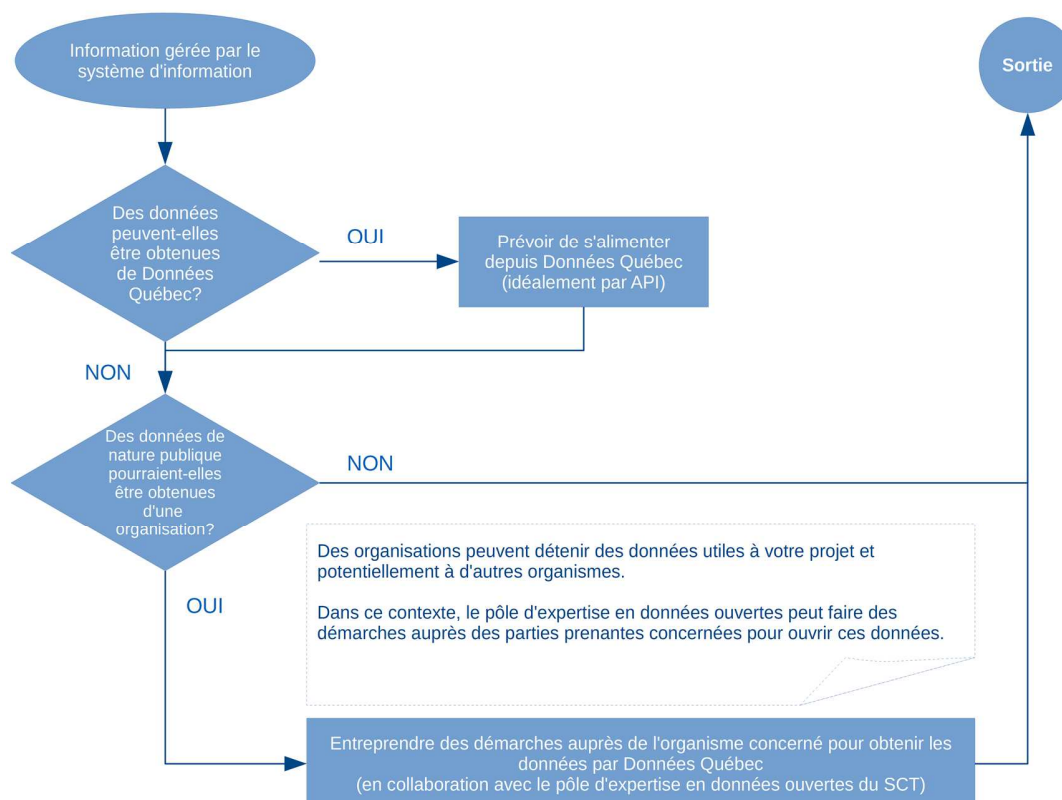
- connaître vos besoins d'affaires;
- lister les informations gérées par votre système d'information.

-

-

2.1 Sous-arbre (volet intrant)

Figure 1 : Sous-arbre sur les données ouvertes (volet intrant)



2.2 Des données peuvent-elles être obtenues de Données Québec?

En s'appuyant sur la liste des informations qui seront gérées par le système d'information, est-ce que certaines données nécessaires sont disponibles en données ouvertes sur Données Québec?

À titre de rappel, les données ouvertes ne contiennent que des informations publiques. Ainsi, aucun renseignement personnel ne sera offert sur Données Québec.

Si des données correspondent aux besoins ciblés, elles peuvent être téléchargées manuellement ou être obtenues à partir de l'[API](#) offerte par Données Québec.

Exemples de données ouvertes utilisées en intrant

Pour appuyer les électeurs à trouver leur circonscription, le Directeur général des élections du Québec (DGEQ) met à la disposition de la population un [outil de recherche](#) réservé à cet effet.

Les utilisateurs peuvent effectuer une recherche par :

- code postal;
- municipalité;
- circonscription (la délimitation de celle-ci s'affichera sur une carte).

Pour assurer une intégrité des données, l'outil de recherche s'appuie sur le [répertoire des municipalités](#) diffusé sur Données Québec.

-

2.3 Des données de nature publique pourraient-elles être obtenues d'une organisation?

Après quelques recherches, il est possible que certaines données publiques ne soient présentement pas offertes en données ouvertes. Bien que de nouvelles données soient ajoutées régulièrement sur le portail Données Québec, la diffusion de données ouvertes est un exercice en continu.

Dans ce cas, il est proposé d'entreprendre des démarches, en collaboration avec le pôle d'expertise en données ouvertes du SCT (pilote@donneesquebec.ca), auprès de l'organisme concerné pour obtenir les données requises de données ouvertes.

2.4 Sous-arbre (volet extrant)

La mesure 32 de la [Stratégie gouvernementale en TI](#) engage l'État à « fournir au public des données ouvertes de qualité qui répondent aux attentes ».

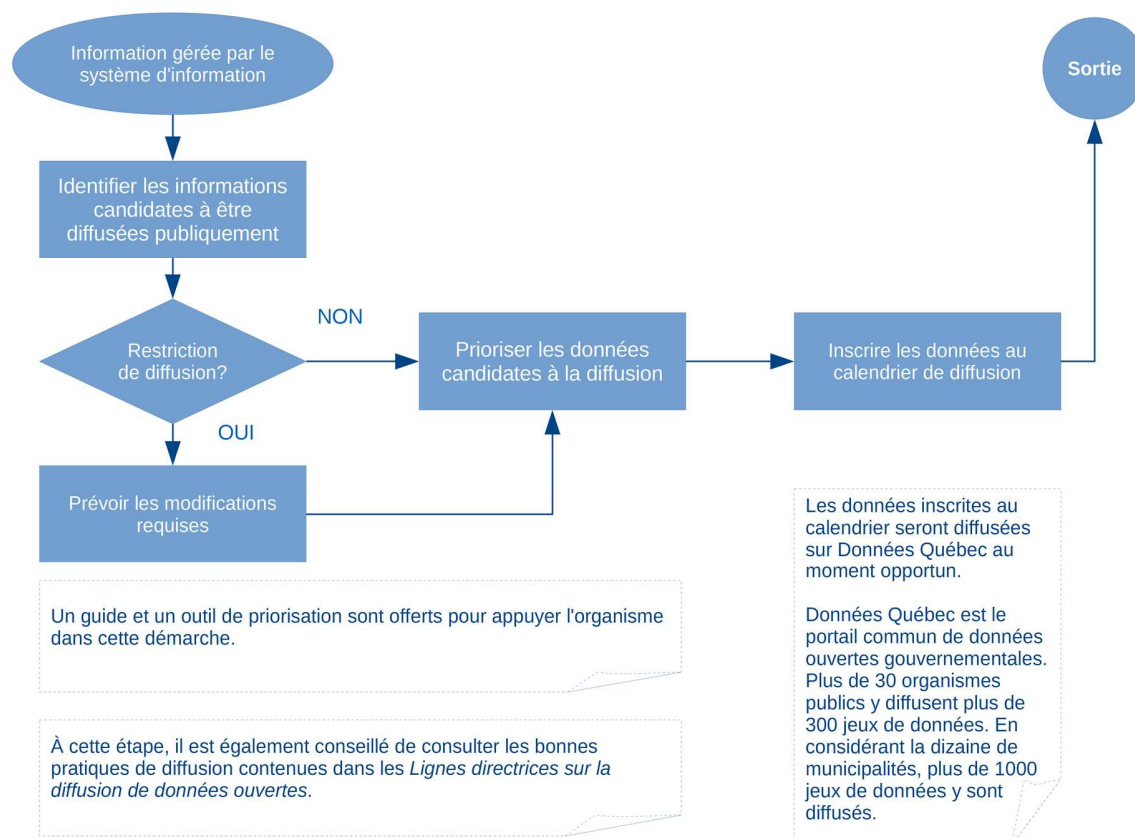
Pour soutenir cette mesure, trois documents et un outil ont été publiés.

- [Guide de priorisation et de diffusion des données ouvertes](#) : accompagner les organisations dans l'identification, la priorisation et la mise en ligne de données ouvertes.
- [Formulaire d'évaluation des jeux potentiels](#) : cet outil accompagne les organisations dans l'étape de priorisation et permet de déterminer les jeux qui feront partie du calendrier de diffusion.

- [Lignes directrices sur les données ouvertes](#) : ce document vient appuyer la préparation de jeux de données en présentant les normes techniques auxquelles les données doivent correspondre afin qu'elles soient déposées sur le portail Données Québec.

Cette section du sous-arbre offre une vue d'ensemble du processus d'identification et de priorisation de données ouvertes pouvant être diffusées de façon proactive par l'organisme. Après que ce sous-arbre ait été parcouru, il sera possible d'inscrire les données candidates au calendrier de diffusion de votre organisme, ce qui permettra de procéder à la diffusion au moment opportun.

Figure 2 : Sous-arbre sur les données ouvertes (volet extrant)



Répondant en gouvernement ouvert de votre organisme

Pour assurer une pleine cohérence au sein de votre organisme pour la diffusion de données ouvertes, il est conseillé de prendre contact avec le répondant en gouvernement ouvert de votre organisme. Au besoin, celui-ci pourra vous appuyer dans cette démarche.

2.5 Identifier les informations candidates à être diffusées publiquement

Cette étape vise à identifier les données candidates à diffuser. En s'appuyant sur la liste des informations gérées par le système, recenser les informations correspondant aux éléments ci-dessous :

- **Les documents énumérés à l'article 4 du Règlement sur la diffusion de l'information et sur la protection des renseignements personnels (PRP), notamment ceux évoqués aux alinéas 6 et 7**
 - o Est-ce que le système est un registre? Produira-t-il des statistiques?
- **Les demandes récurrentes d'accès à l'information**
 - o Est-ce que des informations gérées par le système pourraient faire l'objet de demandes récurrentes?
- **Les données géomatiques**
 - o Est-ce que des données géomatiques seront gérées par votre système?
- **Les autres données partagées auprès des partenaires**
 - o Est-ce que certaines données gérées par votre système devront être diffusées auprès de partenaires (internes ou externes à l'administration publique)?

À partir de la liste dressée, il est possible de regrouper les informations par thème en vue de créer un jeu de données complet et cohérent. Voici quelques exemples de regroupements :

- données concernant des infrastructures (ex. : [attraits touristiques au Québec](#));
- registre d'information (ex. : [répertoire des terrains contaminés](#));
- bilan d'opérations (ex. : [rapports d'accidents](#)).

Pour chacun des jeux de données identifiés, fournir une brève description des principaux champs.

2.6 Restrictions de diffusions?

Pour chacun des jeux de données identifiés à l'étape précédente, une première analyse en matière de restrictions de diffusion est à effectuer :

- **Préciser les enjeux liés à la protection des renseignements personnels (PRP)**
- Déterminez l'information nominative en vue de l'anonymiser ou de la retirer au moment de la diffusion.
- **Recenser les renseignements personnels à caractère public en vue de les retirer**
- Certains renseignements personnels sont déclarés publics par diverses lois dans un objectif précis. Ces renseignements, compte tenu du cadre légal actuel, ne peuvent être inclus dans des jeux de données ouvertes. À cette étape, assurez-vous de les recenser en vue de les retirer avant la diffusion du jeu de données.
- **Recenser les renseignements confidentiels en vue de les retirer**
- Précisez les modifications requises au jeu de données pour qu'il soit possible de le diffuser. Au besoin, à l'aide de la démarche du [Guide de catégorisation de l'information](#), vérifiez qu'une catégorisation de l'information du jeu potentiel permettrait de le diffuser en données ouvertes.
- **Vérifier que l'organisme détient les droits d'auteur**
- Assurez-vous que votre organisme détient bien les droits d'auteur (patrimoniaux et moraux) relatifs aux données que vous prévoyez diffuser. À terme, cette étape vous permettra de vérifier que la licence Creative Commons, désignée par le gouvernement du Québec, peut être appliquée à vos jeux de données potentiels.

Il est possible de consulter une version plus détaillée de cette étape à la section **4.1 Identification des sources d'information** et **4.2 Description préliminaire des jeux de données potentiels** du *Guide de priorisation et de diffusion des données ouvertes*.

2.7 Prioriser les données candidates à la diffusion

L'étape de priorisation permet de cibler les jeux qui feront partie du calendrier de diffusion. Cette étape est réalisée à partir [du formulaire d'évaluation des jeux potentiels](#).

Arrimage avec le répondant en gouvernement ouvert de votre organisme

Il se peut qu'une priorisation à l'échelle de votre organisme soit déjà en cours de réalisation. Pour assurer une pleine cohérence au sein de votre organisme, il est proposé que cette étape soit réalisée en collaboration avec le répondant en gouvernement ouvert.

Ce questionnaire à choix multiples donne des résultats sur deux axes, soit la « **valeur du jeu** » et la « **simplicité de diffusion** ». Les réponses au formulaire, sans être parfaitement exactes, doivent être précises. L'évaluation de l'ensemble des critères de priorisation devrait être réalisée en moins d'une heure par jeu. Les jeux qui auront obtenu davantage de points pour les critères « valeur du jeu » et « simplicité de diffusion » devront normalement être priorisés dans le calendrier de diffusion.

Il est possible de consulter une version plus détaillée de cette étape à la section **4.3 Priorisation des jeux potentiels** du Guide de priorisation et de diffusion des données ouvertes.

2.8 Inscrire les données au calendrier de diffusion

Après la réalisation du questionnaire, un tableau de bord sera dressé automatiquement par l'outil de priorisation en fonction des points obtenus.

À partir de ce tableau de bord, il est possible de communiquer avec le répondant en gouvernement ouvert de votre organisme afin d'inscrire les jeux prioritaires au calendrier de diffusion de votre organisme.

