



PROJET DE NOTE STRATÉGIQUE POUR UNE STRATÉGIE FÉDÉRALE OPEN DATA



L'Open Data est un moteur pour l'innovation, la croissance économique, la transparence et la participation. Les avantages économiques cumulés découlant de la mise à disposition de données publiques peuvent s'élever à 40 milliards d'euros par an dans l'UE. Ces données se prêtent tout à fait à la réutilisation dans la cadre de nouveaux produits et services et peuvent améliorer l'efficacité des pouvoirs publics. Rendre accessibles les données du secteur public permet d'impliquer davantage les citoyens dans la vie politique et sociale et d'apporter une contribution active à des domaines politiques tels que l'environnement, la mobilité, l'économie, etc.

L'accord gouvernemental stipule que le gouvernement s'attèlera au cours de cette législature à l'ouverture de ses données publiques (politique « Open Data »).

Les services publics mettent ainsi à disposition des citoyens, des chercheurs, des entreprises et des pouvoirs publics à des fins de réutilisation des données dont ils disposent déjà dans le cadre de leurs missions.

Il s'agit plus précisément de données :

- que les services publics ont collectées dans le cadre de leurs marchés publics ;
- qui ne contiennent pas (plus) d'informations sensibles sur le plan de la vie privée et qui ne sont pas protégées par des droits de propriété intellectuelle ;
- qui sont publiées dans un format facile à utiliser de façon automatique ;
- qui peuvent être réutilisées à des fins commerciales et non commerciales.

Il ne s'agit pas uniquement de mettre à disposition des données publiques, mais aussi de construire un environnement ("community") autour de ces données. Les entreprises peuvent en effet utiliser certaines données et les enrichir encore avec leur propres informations pour commercialiser de nouveaux produits et services. Les citoyens bénéficient de pouvoirs publics plus transparents et les services publics peuvent aider à améliorer la qualité de données et services publics existants.

1. Exposé du dossier

Ces dernières années, un écosystème « Open Data » s'est développé au niveau mondial. À la suite du Royaume-Uni et des États-Unis, bon nombre de pays (dont les Pays-Bas, la France, l'Allemagne et le Danemark), de régions et d'administrations locales ont commencé à mettre leurs données à disposition du public pour toute forme d'utilisation. La Commission européenne donne aussi de plus en plus souvent accès à ses propres données et a déjà créé un portail pan-européen.

1.1 EUROPE

Au niveau européen, la Commission européenne a fait connaître sa stratégie d'Open Data en décembre 2011, avec pour objectif de faire de l'Open Data la règle dans les États membres de l'Union européenne. La Stratégie Open Data européenne est considérée comme un pan important de l'Agenda numérique européen, misant sur les technologies de l'information et de la communication pour réaliser les objectifs 2020. L'Agenda numérique entend faciliter au maximum l'Open Data et la réutilisation d'informations publiques afin d'aboutir à un marché unique numérique et donner accès aux données publiques.

Il y a déjà longtemps que l'Union européenne vise une plus grande disponibilité des données dont dispose le secteur public et cet objectif a débouché, en 2003, sur la directive concernant la réutilisation des informations du secteur public (directive PSI). Elle avait pour but de favoriser l'émergence d'un marché européen des informations du secteur public via une harmonisation de la réglementation et la création d'un level playing field sur ce marché. En 2011, une révision de cette directive s'est toutefois avérée nécessaire car des évaluations avaient mis en évidence la nécessité d'une réglementation plus contraignante pour parvenir à un véritable marché européen des informations du secteur public. La Commission européenne a lié cette révision de la directive PSI au concept plus populaire d'Open Data, qui faisait aussi l'objet d'un vaste mouvement depuis quelques années.

Cette directive PSI revue de 2013 doit être transposée en droit belge avant le 18 juillet 2015. Pour le gouvernement belge, c'est l'occasion de se doter d'une stratégie fédérale « Open Data », comme il en existe déjà dans bon nombre de pays, de régions et de villes.

1.2 BELGIË

Les régions et quelques villes ont déjà pris plusieurs initiatives en la matière. Le gouvernement flamand mais aussi plusieurs villes belges disposent aujourd'hui d'un portail où le public peut accéder à des informations publiques à des fins de réutilisation. Au niveau fédéral, un mouvement de rattrapage est nécessaire.

À la lumière des comparatifs internationaux, la Belgique enregistre au niveau fédéral un moins bon score que la moyenne pour plusieurs raisons :

- la non-disponibilité comme données ouvertes d'informations importantes (comme les informations de voyage sur les transports en commun, les aperçus détaillés des dépenses...) : ces informations ne sont pas publiées ou ne sont pas disponibles dans un format facilement exploitable ;
- l'absence de licence ou l'existence d'une licence trop limitée (par exemple l'interdiction de l'utilisation commerciale) ;
- l'actualisation non ponctuelle ou non régulière des données.

La Belgique enregistre un score de 360 points sur un maximum de 700 en matière d'Open Data. Elle se classe ainsi à la 16e place parmi les 28 pays de l'UE dans l'Indice de la Commission européenne relatif à l'économie et à la société numériques (« Digital Economy and Society Index »).

2 Qu'est-ce que l'Open Data?

Il s'agit de données qui peuvent être librement utilisées, réutilisées et rediffusées par tous à n'importe quelle fin.

Dans ce cadre, on veille à limiter le plus possible les restrictions de réutilisation, notamment à l'aide de normes ouvertes et de licences ouvertes.

L'essence même de l'Open Data est que les utilisateurs puissent trouver facilement les données, qu'elles soient réutilisables et que l'on n'impose pas inutilement des restrictions d'utilisation ou des conditions à remplir avant l'utilisation des données. Ce concept d'ouverture et cette définition sont cruciaux pour l'interopérabilité. L'interopérabilité désigne la capacité de plusieurs systèmes et organisations à coopérer. Dans ce cas, il s'agit de la possibilité de combiner plusieurs ensembles de données (datasets). L'interopérabilité améliorera significativement la possibilité de combiner plusieurs datasets et permettra de ce fait de développer plus de produits et de services de meilleure qualité.

3 Objectif : des données cinq étoiles

Pour encourager les acteurs qui publient des données ouvertes à mettre à disposition leurs données dans un format aussi réutilisable que possible, Tim Berners-Lee a proposé un modèle à cinq étoiles, les étoiles étant attribuées de la façon suivante :

- ★ L'information est disponible en ligne dans n'importe quel format.
- ★★ L'information est mise en ligne dans un format structuré, qui se prête à une réutilisation automatique (comme Excel au lieu d'une image d'un tableau).
- ★★★ L'information est mise en ligne dans un format de fichier ouvert (comme CSV au lieu d'Excel).
- ★★★★ Pour identifier les choses dans les données, des URI (Unique Resource Identifiers ou adresses web) sont utilisées de manière à ce que d'autres puissent facilement faire référence aux données.
- ★★★★★ Des liens sont établis vers des URI d'autres sources de données pour contextualiser les données.

4. Quel sont les avantages de l'Open Data?

4.1 POUR LES POUVOIRS PUBLICS

- **Économie financière** : étant donné qu'une organisation publique individuelle dotée d'une approche « Open Data » organise autrement ses échanges d'informations, d'autres institutions publiques peuvent aussi en récolter les fruits (financiers).
- **Augmentation de la qualité des données** : l'Open Data est bénéfique pour la qualité des données, car l'organisation est encouragée à améliorer la qualité de ses propres données avant de les rendre publiques ou car des mécanismes de feedback sont mis en place.
- **Renforcement de la légitimité** : la légitimité de l'organisation est renforcée puisque les informations dont l'organisation a la gestion sont utilisées plus souvent. Son existence a plus souvent un impact sur la société.
- **Ouverture accrue de l'organisation sur l'extérieur** : en optant pour une stratégie "Open Data", une organisation est incitée à regarder vers l'extérieur et à collaborer davantage avec des partenaires.
- **Avantages de normalisation** : si le nombre de données exploitées selon des normes (ouvertes) augmente, d'autres acteurs publics peuvent aussi les réutiliser plus facilement, ce qui accroît l'interopérabilité des pouvoirs publics.

4.2 POUR L'ÉCONOMIE

Les bénéfices économiques totaux qui peuvent découler de la mise à disposition de données publiques peuvent s'élever, selon la Commission européenne, à 40 milliards d'euros par an. Le potentiel économique ne réside pas seulement dans la réutilisation d'informations du secteur public pour de nouveaux produits et services, mais aussi dans des gains d'efficacité dans les services publics. Pour la Belgique, le bénéfice net potentiel issu de l'ouverture de données publiques se chiffrerait à 900 millions d'euros, selon les estimations de la fédération sectorielle de l'industrie technologique Agoria.

4.3 POUR LES CITOYENS ET LA DÉMOCRATIE

Il ne faut pas non plus sous-estimer l'importance sociétale de l'ouverture des données. La transparence des pouvoirs publics est décisive pour la participation citoyenne. Dans une vraie démocratie, les citoyens prennent part au débat sociétal sur la base d'informations correctes qu'ils peuvent utiliser pour analyser des choix politiques, poser des affirmations et examiner des faits. Comme les pouvoirs publics ne savent pas toujours ce qui est pertinent pour le citoyen, les données brutes doivent aussi être mises à disposition. Organisations, citoyens, entreprises, etc. peuvent ainsi s'atteler à générer eux-mêmes des informations pertinentes et à participer activement au développement de solutions créatives.

5. Grandes orientations stratégiques en vue d'une Stratégie fédérale Open Data

Les pouvoirs publics mèneront désormais une politique ouverte par défaut ("open by default") basée sur les principes suivants.

1. COMPLY OR EXPLAIN

Toutes les données qui sont récoltées dans le cadre de missions de services publics seront considérées comme ensemble potentiel de données ouvertes (« open dataset ». Ces données seront mises à disposition pour réutilisation, sauf motif juridique contraire (principe "comply or explain").

2. MISE À DISPOSITION PROACTIVE

D'ici 2020, la mise à disposition des données sera proactive. Lors de la phase de transition, elle pourra se faire sur simple requête de l'utilisateur si nécessaire. L'on imposera pas de nouvelles formalités (comme l'enregistrement obligatoire) aux entreprises, citoyens... qui veulent consulter ou réutiliser les données.

3. OBJECTIF : DES DONNÉES CINQ ÉTOILES

Les services publics tendent à publier des données cinq étoiles pour permettre l'interopérabilité et favoriser la réutilisation. Les informations restent toutefois disponibles dans un format facilement accessible pour l'utilisateur non-technique.

4. RENFORCEMENT DES SOURCES DE DONNÉES AUTHENTIQUES

Les données ouvertes sont autant que faire se peut issues des sources de données fédérales authentiques, c'est-à-dire de sources gérées par le service public compétent. Cette approche doit déboucher sur un ensemble de données publiques fiables et qualitatives, et limiter le risque d'erreurs ou de mauvaise utilisation.

5. MODÈLES DE LICENCE SIMPLES

Des modèles de licence simples et standardisés sont utilisés. La "licence ouverte" (licence CC0) qui permet la réutilisation gratuite des données à des fins commerciales et non commerciales est d'application là où c'est possible. Si ce n'est pas le cas, cela doit être explicitement motivé par le service public concerné (principe 'comply or explain').

6. GESTION ET CONTINUITÉ

Chaque service public reste responsable de la publication et de la gestion de ses propres données et de ses métadonnées. La continuité de la disponibilité des données est la règle. Si le service public arrête de publier certaines données, il doit motiver sa décision. Les métadonnées sont décrites à l'aide de la norme européenne (DCAT-AP) concernant les portails de données.

7. TASKFORCE, « OPEN DATA CHAMPIONS » ET STRATÉGIE POUR CHAQUE SERVICE PUBLIC

Une taskforce commune est créée à l'intérieur de Fedict et de l'ASA pour soutenir la mise en œuvre de cette stratégie. Un protocole de coopération régit la coopération entre Fedict et l'ASA. Par ailleurs, tout service public désigne un responsable de l'Open Data (« Open Data champion ») qui fait office de point de contact pour la taskforce et se charge d'élaborer une stratégie « Open Data » pour le service public concerné. Les "Open Data Champions" siègent au comité de transparence prévu par la loi.

8. UN PORTAIL FÉDÉRAL UNIQUE

Un portail fédéral unique renvoie vers les open datasets des différents services publics et vers les portails des entités fédérées, des autorités locales et le portail pan-européen « Open Data ». Ce site internet peut être utilisé par des développeurs mais est aussi accessible à un plus large public.

9. PROMOTION ET DIALOGUE DANS UNE COMMUNAUTÉ OPEN DATA

Les pouvoirs publics informeront le public quant à la disponibilité des données et feront la promotion de leur réutilisation. Des initiatives sont prises pour instaurer un dialogue dans la communauté des réutilisateurs (potentiels) et créer des interactions avec les services publics concernés. Un mécanisme de feedback structuré sera facilité pour améliorer la qualité de la source authentique de données et pour que les citoyens et les entreprises puissent suggérer d'éventuelles améliorations.

10. « INVENTAIRE OPEN DATA » ET CALCULUS

La taskforce au sein de Fedict et de l'ASA dresse l'inventaire des sources de données présentes dans les services publics fédéraux et les complète avec les données suivantes :

- revenus issus de la source de données ;
- coût de l'ouverture ;
- nombre d'étoiles attendu à l'ouverture ;
- statut et plan d'action pour la poursuite de l'exploitation de la source de données.

L'inventaire est publié et régulièrement mis à jour.

La taskforce examine la possibilité d'élaborer une méthode de mesure (calculus) permettant de quantifier la valeur ajoutée que représente l'ouverture de sources de données spécifiques.

11. FONCTIONNEMENT AXÉ SUR LA DEMANDE

Dans le cadre de l'ouverture de données, les services publics donneront la priorité aux datasets demandés par les citoyens et les entreprises via leur propre site internet, via le portail fédéral ou qui sont utilisés dans les comparatifs internationaux comme le Baromètre Open Data. La taskforce peut à cet égard donner un avis sur le fond et un avis technique, mais l'ouverture en elle-même est réalisée par les services publics.

12. APPUI DES SERVICES PUBLICS

La taskforce créée au sein de Fedict et de l'ASA accompagnera les services publics dans le cadre du passage à une politique ouverte par défaut.

13. OPEN DATA ET MARCHÉS PUBLICS

Les services publics veillent contractuellement à ce que lors de nouveaux marchés publics où des données sont recueillies pour réaliser un produit ou un service, ils obtiennent les droits nécessaires pour pouvoir publier les données sous une licence ouverte. Ils veillent aussi contractuellement à ce que les données soient exploitées et structurées de manière à faciliter plus tard la réutilisation et à l'utilisation de normes ouvertes.

14. OPEN DATA DANS LES CONTRATS DE GESTION

Dans les contrats de gestion avec les services publics, les organismes et les entreprises publics se verront imposer les mêmes principes et obligations, tels qu'exposés dans cette stratégie.

2.4 OPEN DATA ET PROTECTION DE LA VIE PRIVÉE

L'Open Data implique la mise à disposition de données publiques à caractère non-personnel. Néanmoins, des défis se posent en termes de protection de la vie privée : la contribution de l'Open Data à l'environnement Big Data, dans lequel la protection des données à caractère personnel représente un défi permanent.

Le pouvoir fédéral a donc recours à une approche équilibrée visant à garantir une protection optimale des données à caractère personnel.

- La loi stipule que les données publiques à caractère personnel ne peuvent pas être mises à disposition à des fins de réutilisation sans autorisation explicite de la Commission de Protection de la Vie privée.
- La législation sur la vie privée est le fil conducteur dans le cadre du processus de sélection des datasets mis à disposition pour réutilisation.
- Un comité sectoriel PSI est créé au sein de la Commission de Protection de la Vie privée qui peut être consulté par les services publics lorsqu'ils appliquent des techniques d'anonymisation ou de manière plus générale, donner des avis sur l'impact potentiel de leur stratégie "Open Data" sur la protection de la vie privée. Si nécessaire, il formule des recommandations à l'égard des services publics sur les mesures à prendre pour protéger la vie privée dans le cadre de l'ouverture de leurs données.