OPEN DATA

Introduction :

L'Open Data a pour but de mettre à disposition des données publiques et libres de droits. Ces données peuvent donc être exploitées et réutilisées librement. Les critères essentiels de l'Open Data sont la disponibilité, la réutilisation, la distribution et la participation universelle.

Définitions des critères

Disponibilité et accès : Les données doivent être pleinement accessibles, moyennant un coût de reproduction raisonnable. De préférence, elles se téléchargent sur Internet. La forme doit être confortable et modifiable.

Réutilisation et redistribution : Les données doivent être fournies sous des conditions permettant la réutilisation et la redistribution, incluant le mélange avec d'autres ensembles de données.

Participation universelle : Tout le monde doit être en mesure d'utiliser, de réutiliser et de redistribuer les données. Il ne doit y avoir aucune discrimination concernant les fins d'utilisation, ou contre des personnes ou des groupes. Par exemple, des restrictions non commerciales qui empêchent l'utilisation commerciale, ou les restrictions d'usage à certains secteurs, ne sont pas compatibles avec l'Open Data.

L'interopérabilité

Définition : L'interopérabilité désigne la capacité de différentes entreprises ou systèmes à travailler ensemble. En l'occurrence, l'interopérabilité est la capacité de mélanger différents ensembles de données.

Grâce à l'interopérabilité, la création de systèmes complexes et vastes est possible et facilitée. En effet, dans le cas des données, l'interopérabilité est essentielle pour dégager des bénéfices de l'ouverture et permettre un développement de produits et de services plus facile, en plus grande quantité et dans une meilleure qualité.

Pour qu'elle soit rendue possible, il faut que les données soient partagées sous un même langage ou qu'il existe un élément programmatique pour lire ces données. Afin que les données soient partagées et comprises de tous, le W3C recommande l'application des standards de l'Open Data (langage Web Sémantique). Si ils sont utilisés, le créateur des données doit certifier la provenance, indiquer les métadonnées liées, garantir la qualité de l'information et son auteur.

Provenance des données

L'Open Data est applicable dans tous les sujets et à n'importe quelles sources. Tout le monde peut proposer des données sous licence ouverte pour un usage gratuit afin que toute personne puisse l'utiliser librement. La majorité des données en Open Data proviennent du gouvernement et du secteur public. Les entreprises, les startups, les ONG ou les individus peuvent aussi proposer de l'Open Data. Ce sont donc des données non personnelles, qui ne comportent aucunes informations sur des individus.

Open Data / Big Data

Domaine d'utilisation

L'Open Data est utilisé dans une multitude de domaines et sujets différents. C'est également le cas du Big Data. L'un rend les données accessibles au public, libres de droit et d'utilisation, l'autre utilise des données recueillies en masse puis fait des liens entre ces données et les traite pour arriver à un résultat. Le point commun entre ces deux systèmes, c'est leurs domaines d'utilisation très vastes et applicables à énormément de sujets.

Exemple : -Transport : Le Big Data, grâce aux données collectées en masse à partir de la géo localisation (GPS), met en évidence l'évolution du trafic et donc la position des bouchons et autres problèmes. Le secteur du Transport utilise ensuite l'Open Data afin de partager les informations sur le trafic routier (ralentissements, bouchons).

L'impact du RGPD

Pour rappel, le RGPD est un règlement qui protège les données à caractères personnels des citoyens européens. Ce règlement, depuis sa mise en vigueur, pose des problèmes aussi bien pour l'Open Data que le Big Data.

Pour l'Open Data, cela pose un problème majeur car les deux projets semblent incompatibles au premier abord. En effet, avant l'application du RGPD, le principe de l'Open Data étant de partager des données, la notion de vie privée n'était pas toujours respectée ce qui posait des problèmes de confidentialité. Depuis l'entrée en vigueur du RGPD, toutes les données partagées par l'Open Data doivent être anonymes et prennent en compte la vie privée.

Pour le Big Data, le problème est similaire à celui de l'Open Data, avant le RGPD, la collecte des données personnelles était moins réglementée de ce fait, le Big Data collectait librement les données personnelles. L'apparition du Big Data à mit fin à cette libre collecte.

Dans les deux cas, l'autorisation des utilisateurs est nécessaire pour recueillir leurs données personnelles.

Intérêts

Intérêts financiers : L'Open Data n'a pas vocation à faire du profit contrairement au Big Data qui a un coût et permet aux entreprises spécialisées dans le Big Data de gagner de l'argent en revendant les données à d'autres acteurs du marché (annonceurs, entreprises...).

Intérêts informatifs : L'Open Data permet la mise à disposition de données destinées au public dans des domaines divers et variés, ce qui permet à tous de les consulter et les utiliser. Le Big Data permet aussi d'obtenir des informations mais ces dernières sont réservées aux demandeurs (qui ont bien évidemment payé), ces données leurs sont réservées.

Conclusion

L'Open Data et le Big Data n'ont pas toujours la même vocation, mais récoltent tous deux des données en masse. La finalité n'est pas identique mais le principe est le même.

Glossaire

Élément programmatique : élément d'un Applicatif / Programme

W3C (World Wide Web Consortium) : Organisme de standardisation des technologies du web

Langage Web Sémantique : extension du web standardisée encourageant l'utilisation de formats de données et protocoles commun à tous.

From:
https://wiki.sio.bts/ - **WIKI SIO : DEPUIS 2017**

Permanent link:
https://wiki.sio.bts/doku.php?id=1d

Last update: **2020/07/26 16:27**