## À quel point mes données sont-elles utilisables?

Évaluer l’utilité et l’exploitabilité de l’open data dépend du domaine et du contenu. Pour faciliter ce processus, il existe un certain nombre de directives de pratiques optimales qu’il est recommandé d’appliquer, que l’on publie ou que l’on utilise les données.

Au cours de ce module, nous étudierons les 5 étoiles des ‘Données Ouvertes et Liées’ (en anglais: LOD ou ‘Linked Open Data’, terme généralement utilisé même en français) et tenterons d’expliquer comment les utiliser pour mesurer la facilité d’utilisation technique des données.

##### Dans ce module, nous allons explorer les points suivants :

* Quelles sont les cinq étoiles du ‘Linked Open Data’
* Les trois premières étoiles
* Comment reconnaître les étoiles dans les données

## Les 5 étoiles de l’open data

Les directives des 5 étoiles du Linked Open Data permettent de mesurer l’intégration des données sur le Web.

Elles examinent l’accessibilité et la facilité d’utilisation technique d’une série de données, allant de 1 étoile (disponibilité en ligne) à 5 étoiles (appartenance à un web des données). Les étoiles sont attribuées de manière séquentielle; on ne peut pas attribuer une étoile supérieure si les exigences de la précédente ne sont pas satisfaites.

Ces recommandations ont été développées par Sir Tim Berners-Lee en 2001 et ont depuis été adoptées par ceux qui publient de l’open data à travers le monde, guidant de nombreuses initiatives d’open data comme celles de l’Agence Numérique d’Italie.

[Lien vers l’Agence d’Italie numérique](http://spcdata.digitpa.gov.it/data.html)

Ce module se concentre sur les 3 premières étoiles, qui sont adaptées aux données tabulaires et transactionnelles. Nous étudierons les 2 dernières étoiles en détail dans le [module 13](http://docs.google.com/modul13/).

## Les 3 premières étoiles

Les 3 premières étoiles de l’open data vous permettent d’établir si oui ou non les données sont utilisables.

### 1 étoile - une licence ouverte

La première étoile est octroyée à toutes données que l’on peut qualifier d’ouvertes, c’est-à-dire celles dont le contenu est disponible sous licence ouverte et ce sous n’importe quelle forme.

Indépendamment de la qualité d’un ensemble de données, celles-ci ne peuvent pas bénéficier de la première étoile si elles ne sont pas disponibles sous licence ouverte.

Un fichier PDF disponible sur un site Web sous licence ouverte remplit les conditions nécessaires pour obtenir la première étoile.

### 2 étoiles - une forme réutilisable

La deuxième étoile est octroyée aux données réutilisables facilement. Les directives stipulent que les données doivent être disponible sous ‘un format largement réutilisable et structuré,’ lisible par une machine et compréhensible par l’humain.

La clé pour atteindre les 2 étoiles est de sélectionner le format le plus réutilisable. Dans certains cas, il se peut que le format le plus réutilisable soit un format propriétaire ou fermé, comme c’est le cas pour un fichier Excel ou Numbers.

Il est préférable de rendre ses données disponibles sous n’importe quel format, quel qu’il soit, plutôt que de ne pas les rendre disponible.

### 3 étoiles - une forme ouverte

Les directives pour la troisième étoile de l’open data stipulent que les données doivent être disponibles sous un format structuré et lisible par une machine qui n’est pas limité à un pack logiciel en particulier.

Un exemple d’un ensemble de données auquel on accorderait 3 étoiles serait un fichier CSV avec une licence ouverte appropriée.

L’entreprise d’énergie italienne Enel présente un bon exemple d’une série de données 3 étoiles qui affiche clairement son classement et le nombre d’étoiles octroyées.

[Lien vers la série de données de 3 étoiles d’Enel](http://data.enel.com/node/4679)

## Exploiter l’utilité de l’open data

Chaque étoile acquise représente une exploitabilité améliorée de l’ensemble de données. Utilisez la liste ci-dessous afin d’évaluer combien d’étoiles accorder à un ensemble de données.

### Étoile 1er - vérifier la licence

Trouver la licence pour un ensemble de données ouvertes peut se révéler l’une des étapes les plus difficiles. Consultez les endroits suivants.

Une licence ouverte autorise explicitement l’utilisation des données à des fins commerciales ainsi que non-commerciales.

Essayez de la retrouver dans les rubriques suivantes :

* Les métadonnées
* Les conditions générales du site ou de la page Web
* Cliquez sur le fichier de données, il y a peut-être une fenêtre popup
* Regardez dans le fichier lui-même

Comme chaque site Web a une approche différente en ce qui concerne l’affichage des informations relatives à la licence, il se peut que vous mettiez un certain temps avant de trouver l’information requise.Étoile 2eme - vérifier l’utilitéQuand vous trouvez de l’open data sur le Web, il y a plusieurs manières d’en évaluer l’utilité. Considérez les questions suivantes :

### Pensez-vous que le format sous lequel les données sont publiées soit le plus facile à utiliser ?

* + Le fichier est-il identique à celui utilisé par ceux qui publient les données ?
  + Y a-t-il des informations supplémentaires ajoutées par un humain ?

Étoile 3eme - vérifier l’ouverture

### Le fichier conserve-t-il sa fonctionnalité complète quel que soit le programme utilisé pour l’ouvrir ?

* Lorsqu’on vérifie le format du fichier (par exemple CSV) en ligne, apprend-on qu’il est nécessaire de télécharger un programme spécifique pour l’ouvrir ?

Le format CSV est le format ouvert le plus populaire et il est facile à identifier. Le format CSV remplit les conditions pour obtenir 3 étoiles tant que les conditions relatives aux deux étoiles précédentes sont remplies. Le portail d’open data bulgare montre de bons exemples de séries de données en format CSV 3 étoiles, où la licence et le format de fichier sont clairement indiqués.

[Lien vers le portal d’open data bulgare](https://opendata.government.bg/dataset/tsentralen-registar-na-yuridicheskite-litsa)

Beaucoup de portails d’open data publient aussi des fichiers sous le format ‘Open Document Format’ (ODF). Bien que leur nom suggère que ces fichiers soient parfaitement ouverts, ils donnent lieu à une certaine confusion lorsque l’on tente de nommer l'outil qui peut ouvrir ces fichiers.

## Êtes-vous prêt(e) à évaluer l’ouverture des données ?

Il y a trois choses simples à prendre en compte lorsque l’on évalue à quel point des données sont ouvertes. Vous en souvenez-vous ?

## Comment évaluer le caractère ouvert d’un ensemble de données grâce aux 5 étoiles ?

Comment procéderiez-vous pour évaluer le caractère ouvert d’un ensemble de données que vous auriez trouvées en ligne ?

* En regardant les critères de chaque étoile individuellement et en attribuant l’étoile la plus élevée atteinte par l’ensemble de données.
* En commençant par les critères de la première étoile et en évaluant de manière séquentielle selon chaque étoile.
* En tentant de deviner

##### Vous avez raison !

Chaque étoile est attribuée de manière séquentielle; on ne peut pas attribuer une étoile supérieure si les exigences de la précédente ne sont pas satisfaites.

##### Êtes vous sûr(e) ?

Chaque étoile est attribuée de manière séquentielle; on ne peut pas attribuer une étoile supérieure si les exigences de la précédente ne sont pas satisfaites.

## Que représente la deuxième étoile ?

La deuxième étoile est octroyée pour...

* Une licence ouverte
* Une forme ouverte
* Une forme réutilisable

##### Vous avez raison !

La deuxième étoile est accordée en fonction de la facilité à réutiliser des données et recommande que les données soient disponibles sous le format le plus réutilisable qui soit, même s’il s’agit d’un format propriétaire.

##### Êtes vous sûr(e) ?

La deuxième étoile est accordée en fonction de la facilité à réutiliser des données et recommande que les données soient disponibles sous le format le plus réutilisable qui soit, même s’il s’agit d’un format propriétaire.

## Évaluer l’open data

Évaluer l’utilité et l’exploitabilité de l’open data dépend du domaine et du contenu. Pour faciliter ce processus, il existe un certain nombre de directives de pratiques optimales qu’il est recommandé d’appliquer, que l’on publie ou que l’on utilise les données.

Les directives des cinq étoiles pour le Linked Open Data permettent de mesurer l’intégration des données sur le Web. Les étoiles sont attribuées de manière séquentielle et aucune étoile ne peut être sautée ou omise.

##### Lors de ce module, nous avons étudié les trois premières étoiles :

* La licence ouverte
* La forme réutilisable
* La forme ouverte

Référez-vous au [module 13](http://docs.google.com/module13/) pour plus de détails sur les deux dernières étoiles.