

Journée AFNOR CN 46 / BnF du vendredi 24 juin 2016
« Bibliothèques, archives et musées à l'heure de l'Open data
Bonnes pratiques, normes et retours d'expérience »

Conclusion

Emmanuelle Bermès, BnF, adjointe aux questions scientifiques et techniques,
direction des Services et des réseaux

La pratique de l'*Open data* suppose plusieurs démarches :

- il faut d'abord choisir une licence, ce qui suppose de **qualifier** les données à ouvrir.
- il faut aussi les **structurer** pour que les utilisateurs sachent comment sont faites les données. Cela suppose des normes.
- il faut ensuite les **gérer** : ouvrir nous engage (en particulier à la mise à jour des données). Cela suppose des choix.
- Enfin, il faut **partager** les données, ce qui recouvre l'idée d'une communauté.

À chacune de ces démarches sont associées de nombreuses questions : celles qui vont suivre, en allant à rebours d'une pensée trop lisse, sont destinées à nous aider à prendre du recul pour aller plus loin dans l'*Open data*.

Qualifier

Avons-nous fait le tour de toutes les données à ouvrir ?

En bibliothèque, nous avançons, mais il existe d'autres données, en particulier documentaires : les données de la recherche, les données d'utilisation des bibliothèques, les données de logs...

Dans tout projet européen, il faudrait une clause d'ouverture des données.

L'enjeu est pour nous de savoir réutiliser, développer l'ensemble de ces données, sans nous limiter à celles que nous considérons actuellement dans nos projets d'*Open data*.

Avons-nous arrêté de câliner¹ nos données ?

Nous pouvons en douter si nous pensons au débat au sein d'Europeana à propos de la licence CC0².

De même, nous sommes réticents à l'idée d'ouvrir des données qui n'ont pas la qualité que nous souhaitons : par exemple, les données des éditeurs sur le site *Nouveautés Éditeurs*³ ne sont pas structurées comme nous le voudrions mais elles sont fraîches, or c'est cette fraîcheur que recherchent précisément nos publics.

Structurer

Le web sémantique est-il si important ?

Il semble que nous pouvons commencer à l'oublier un peu. Il a été essentiel pour structurer les données, pour construire des graphes... Aujourd'hui, il n'est plus quelque chose d'innovant : il est bel et bien entré dans les usages des professionnels, qui en font comme Mr Joudain faisait de la prose. Ainsi, même si elle n'est pas présentée de la sorte, l'API IIF est du web sémantique.

¹ Cf en anglais, « data hug »

² Créative Commons Zéro

³ <http://nouveautes-editeurs.bnf.fr/>

Par ailleurs, le web sémantique nous parle parce qu'il parle des données, mais il ne vient pas du monde de l'informatique.

Ce qu'il faut garder du web sémantique, ce sont les identifiants, les données structurées...

Gérer

Avons-nous acquis les bonnes compétences ?

De même que la télévision n'a pas remplacé le cinéma, le web sémantique n'a pas remplacé les bases de données relationnelles... À l'heure du *Big Data*, il est d'abord important d'être en mesure de faire des liens et de se doter d'un outillage pour assurer d'une part la sauvegarde des supports, et d'autre part l'exploration des contenus (par exemple via ElasticSearch...). C'est là que le dialogue avec les informaticiens est essentiel.

Il est également important de développer notre capacité à mettre en forme les données, à pratiquer le *storytelling*.

Partager

Quelques ingrédients pour un partage réussi :

- commencer par utiliser soi-même ses propres produits pour en tester la qualité et l'intérêt : c'est le principe du dogfooding (« eat your own dog food »)
- penser des communautés horizontales
- fournir du bon café dans les hackathons

La question de la soutenabilité

À ces perspectives, il convient d'ajouter un cinquième point : tout cela coûte cher.

Rappelons-nous la façon dont l'*open source* a contribué à faire baisser la valeur du logiciel lui-même. L'ouverture des données génère des économies.

Allons plus loin : les données ont été comparées au sang par Denis Berthault. La circulation du sang n'apporte rien à notre vie, il la conditionne : il en va de même pour les données.

Ainsi, la question de l'*Open data*, loin de se réduire à des questions techniques, constitue un sujet de société à part entière : à l'heure où, dans le monde, un nouveau clivage imprime de plus en plus sa marque entre ceux qui se referment sur eux et ceux qui veulent s'ouvrir, les données, en circulant, rendent le monde meilleur. S'ouvrir, c'est se rassembler pour peser, favoriser la démocratie, défendre des valeurs, constituer une mémoire et un patrimoine commun. C'est fondamental dans nos sociétés individualistes, qui mettent par exemple en valeur le droit à l'oubli.