

archives et Wikidata

Le travail d'archiviste en Suisse a été en grande partie jusqu'ici un labeur solitaire et artisanal. Est-il destiné à le rester? Face aux défis que pose la maîtrise actuelle de l'information, on peut en douter. Cet article envisage le rôle que pourrait jouer la plateforme Wikidata, et, par ce biais, la technologie des données ouvertes liées, pour renouveler le travail des archivistes.

« *Je rêve d'un âge nouveau de la curiosité. On en a les moyens techniques; le désir est là; les choses à savoir sont infinies; les gens qui peuvent s'employer à ce travail existent. De quoi souffre-t-on? Du trop peu: de canaux étroits, quasi monopolistiques, insuffisants* » (Le philosophe masqué) (1). Le manque de visibilité et d'accessibilité en ligne des descriptions des archives est un fait, mais on dira peut-être un jour à ce propos: « *Et Wikidata est arrivé* ». Comme on peut déjà le dire pour Wikipédia à propos du savoir encyclopédique, tous les ingrédients semblent en effet rassemblés pour que Wikidata, lancé en 2012, devienne le nœud absolument central de notre accès à l'information. On pourrait y voir comme une réponse tardive aux songes du philosophe masqué.

description d'archives

L'intérêt de la plateforme Wikidata vis-à-vis des documents regroupés au sein du portail Wikimedia Commons ayant déjà

été évoqué (2), nous allons pour notre part en examiner son grand intérêt pour la description d'archives en général (3). Cet article s'insère dans un contexte foisonnant d'expérimentations, en Suisse comme ailleurs, autour des données ouvertes liées (« *linked open data* » ou LOD, en anglais), qui semblent offrir des perspectives nouvelles et enthousiasmantes pour nos métiers, comme en atteste, parmi d'autres, le tout récent et important document auxiliaire publié par eCH à ce sujet (4).

Son propos pourrait être résumé ainsi: les méthodes de description sont tributaires des technologies et des outils techniques pour les mettre en œuvre. Or la technologie des données ouvertes liées semble ouvrir un nouveau chapitre en cela. Dès lors, quels outils techniques choisir? Notre article propose d'utiliser massivement le portail mutualisé Wikidata pour toutes les données issues des archives pour lesquelles ce n'est pas illégal.

Mais qu'est-ce que Wikidata? Elle se définit comme « *une base de données libre, collaborative, multilingue et secondaire qui collecte des données structurées pour alimenter [...] les autres projets du mouvement Wikimedia et pour n'importe qui sur la planète* » (5).

Ce qu'il faut en savoir:

- cette base de données collaborative en ligne répond aux standards des données ouvertes liées;
- elle dispose de critères de recevabilité des informations plus larges que Wikipédia. Toute information, du moment où sa source est indiquée, est recevable dans Wikidata (6);
- son contenu vise à alimenter dynamiquement d'autres portails (Wikipédia, Dbpedia, etc.);
- son interface principale (Wikidata.org) n'est pas une interface de consultation mais de saisie.

Wikidata répond aux principes des données ouvertes liées. Elle est donc un outil très actuel. À la pointe de la technologie

contemporaine. Mais ce n'est pas seulement cela, car c'est aussi un projet récent de la Wikimedia Fondation et, en ce sens, le résultat de l'expérience accumulée par celle-ci dans le développement d'outils en ligne d'usage gratuit, collaboratifs, multilingues et utilisant des données aux formats et licences ouverts. En ce sens, si Wikidata est un outil de gestion et publication de données ouvertes liées parmi d'autres, son grand intérêt réside dans sa disponibilité en ligne, sa gratuité, son outillage collaboratif évolué et son code ouvert.

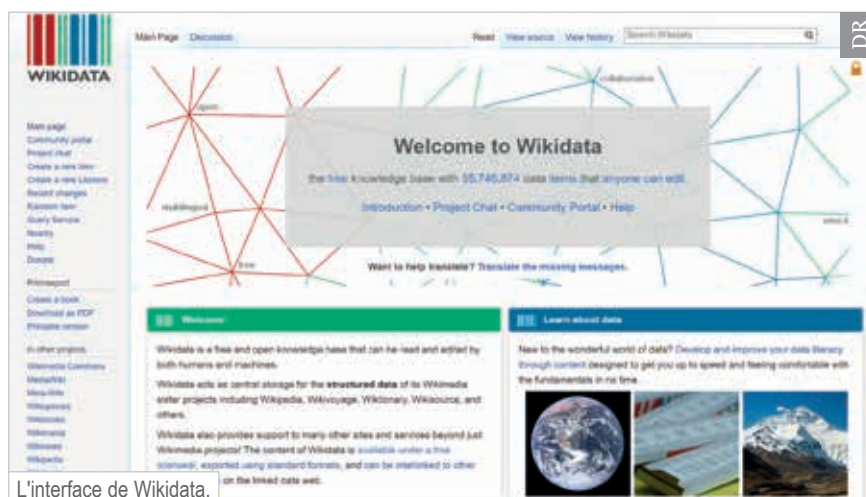
Wikidata et coconstruction

Ce que Wikidata peut irrémédiablement changer, c'est la mutualisation des efforts des archivistes vis-à-vis de la description. Jusqu'à présent, ce que les archivistes avaient espéré et tenté, c'était de pouvoir échanger et interroger leurs données respectives. Chacun restait derrière ses portes closes, mais retrouverait la fenêtre de son guichet de consultation de données. Le projet envisageable avec Wikidata est inverse: toutes et tous travaillent de façon collaborative sur les mêmes données. Les données partagées et ouvertes. Toutes et tous travaillent ensemble à bâtir une base de connaissance commune. On rejoint ici complètement Beat Estermann qui propose, dans une vision plus large que le projet, que Wikidata devienne « *a central hub for data integration, data enhancement, and data management in the heritage domain* » (7).

Selon notre hypothèse, toutes les données qui peuvent légalement l'être devraient être transférées sur Wikidata. Et quand nous disons transférées, cela ne veut pas dire établir un lien d'homologie entre des entités, par exemple entre notices d'autorité comme cela peut être le cas au travers du projet MetaGrid (8), mais bien que l'entité référencée soit celle de Wikidata.

« est-il venu le temps que les archives se libèrent de l'autorité des archivistes ? »

Baptiste de Coulon



L'interface de Wikidata.

mutualisation des efforts

Nous y voyons une cascade de conséquences positives :

- cela pousserait les institutions à enrichir les données dans Wikidata et améliorerait exponentiellement la qualité de celles-ci et l'intérêt de cette plateforme pour la recherche;

- cela déplacerait une grande partie de la recherche d'informations vers les différents modes d'accès aux données contenues dans Wikidata, qui offre déjà des moyens de recherche très puissants (par exemple via le langage de requête SparQL) et inédits dans le domaine des archives (9);

- cela inciterait les institutions à investir dans des outils de visualisation des résultats de requêtes issues de l'interrogation de Wikidata.

Ces résultats sont perçus comme positifs, car cela mutualiserait les efforts autour d'une base de connaissance commune (et des moyens d'y accéder) au lieu de les disperser dans la multiplication des initiatives institutionnelles individuelles pour des résultats mitigés. Cela serait aussi un moyen simple de suivre cette recommandation qui semble tomber sous le sens : « *Public money? Public code!* » (10).

Wikidata et RIC

Enfin, utiliser massivement Wikidata signifie aussi pour les institutions qui conservent des archives s'approprier

facilement et rapidement les recommandations de base contenues dans le dernier projet en date de norme de description archivistique intitulé « *Records in contexts* » (RIC) (11), qui propose justement de fusionner les normes de description existantes en utilisant tout le potentiel des données ouvertes liées.

Le 19^e siècle a vu l'école s'émanciper de sa tutelle religieuse; est-il venu le temps que les archives se libèrent de l'autorité des archivistes? Nous serions tentés de plaider en ce sens. Cela répond à un principe de réalité, les archivistes ne peuvent plus prétendre réaliser seuls la sauvegarde des informations essentielles et par ailleurs, nous souhaiterions qu'ils cessent d'en être les cerbères.

Si le monde est la coconstruction de nos actions, de quel monde voulons-nous en tant qu'archivistes être bâtisseurs? Le mouvement Wikimedia a choisi : il s'est doté d'une feuille de route « *Wikimedia 2030* » (12), grâce à laquelle il veut bâtir une infrastructure essentielle au service du savoir libre et référencé.

En serons-nous partie prenante? Tout porte à croire que les archivistes ont un réel rôle à y jouer. Mais pour y arriver, ils doivent libérer leurs données et se libérer de leur autorité sur elles. Nous avons tenté de proposer des pistes pour que les archivistes participent à ouvrir des temporalités émergentes et des futurs inédits (13).

Alors intéressé-e-s? Poursuivons la réflexion ensemble (14)! ■

Baptiste de Coulon

(1) Michel Foucault cité par Patrick Boucheron « *Écrire l'histoire des futurs du passé* » in Patrick Boucheron, François Hartog, *L'histoire à venir*, Anacharsis, Toulouse, 2018.

(2) Rama Neko, « *Données structurées, la puissance de Wikidata au service de Wikimedia Commons* », arbedo, no 03 (2017)

(3) Pour un aperçu de réflexions très proches mais dans le domaine des bibliothèques, voir l'article de Stacy Allison — Cassin et Dan Scott « *Wikidata: a platform for your library's linked open data* », code4lib journal, issue 40, 4 mai 2018.

(4) → www.ech.ch/alfresco/s/ech/download?nodeid=aee48e8b-2307-44fb-b1d3-ad61ed4432bc

(5) → www.wikidata.org/wiki/Wikidata:Introduction/fr

(6) → www.wikidata.org/wiki/Wikidata:Notability/fr

(7) Wikidata:WikiProject Cultural heritage.

(8) → www.metagrid.ch

(9) → www.wikidata.org/wiki/Wikidata:Data_access/fr

(10) → publiccode.eu

(11) → fr.wikipedia.org/wiki/Records_in_Contexts

(12) → meta.m.wikimedia.org/wiki/Special:MyLanguage/Strategy/Wikimedia_movement/2017/Direction

(13) Nous empruntons ces termes à l'essai stimulant de Jérôme Baschet, « *Défaire la tyrannie du présent; Temporalités émergentes et futurs inédits* », La Découverte, Paris, 2018.

(14) → www.wikidata.org/wiki/Wikidata:WikiProject_Archival_Description