

TWEETS ET PUBLICATIONS INSTAGRAM, UN PATRIMOINE NUMÉRIQUE À PART ENTIÈRE

Publié le 11 janvier 2022



par Camille Stassart

Conserver, gérer et étudier les productions liées au patrimoine culturel et historique belge. Telle est la principale mission de la [KBR, la bibliothèque scientifique de Belgique](#). Alors que plusieurs de ses sources sont hébergées sur Internet, la KBR s'attelle à mettre en place une stratégie durable d'archivage et de préservation du patrimoine national en ligne. L'enjeu est important : ce matériel numérique est, par nature, éphémère. Ne pas l'archiver, c'est prendre le risque qu'il disparaisse un jour définitivement, et donc de ne pas disposer de traces pour les générations futures.

Au travers du [projet BESOCIAL](#), financé par BELSPO, des chercheurs de la KBR, de l'UCLouvain, de l'UNamur et de l'UGent tentent de concevoir une approche pour préserver les contenus des médias sociaux belges. Une opération qui se frotte à de nombreux défis.

De l'importance de conserver le passé d'Internet

« Ce projet est né à la suite du [projet PROMISE](#), qui s'est clôturé en 2019, dont l'objectif était de développer une stratégie pour l'archivage des sites web belges. Avec BESOCIAL, nous nous concentrons sur le contenu diffusé via les réseaux sociaux. Ceux-ci représentent des sources très intéressantes pour divers domaines de recherche comme l'histoire, les sciences politiques, la linguistique, la sociologie, les sciences de l'information, ou encore l'informatique. Leur préservation est essentielle pour aider à mieux comprendre la société d'aujourd'hui et de demain », souligne la

coordinatrice du projet, Sophie Vandepontseele, Directrice des collections contemporaines papiers et numériques à la KBR.

Pour l'heure, il n'existe aucune politique d'archivage pour ce matériel dans notre pays. « On est un peu en retard sur la question », reconnaît Friedel Geeraert, experte en archivage du web à la KBR et participante au projet. « Cela fait quelques années que des initiatives de ce type ont été lancées dans d'autres pays. On peut citer les projets de la Bibliothèque nationale du Luxembourg, de la Bibliothèque nationale de France, ou encore de la Bibliothèque royale du Danemark ».

[La première phase du projet a consisté à analyser une série de programmes d'archivage existants à l'étranger.](#)

Des défis tant techniques que juridiques

« Nous avons constaté que, puisqu'il est impossible de collecter l'ensemble des données présentes sur les médias sociaux, les programmes de notre échantillon recourent à des sélections sur base d'événements spécifiques, de « hashtags », de mots-clés, ou d'une période de temps limitée », informe Fien Messens, chef du projet BESOCIAL. « Par ailleurs, Twitter reste la plate-forme la plus souvent ciblée par les institutions, car ce média social (tout comme Instagram), présente une API (« application programming interface »). »

Cette interface permet de « connecter » le service à un autre logiciel ou service en vue d'échanger des données. Le problème : les API limitent les informations pouvant être recueillies. Comme le précise Friedel Geeraert, « pour Twitter, on ne peut collecter que 3.200 tweets par compte à chaque archivage, du plus récent au plus ancien. Pour Instagram, il y a une limitation générale du débit (qui n'est documentée nulle part). La session expire donc après un certain nombre de demandes. Après chaque dépassement de délai, il faut attendre quelques minutes avant que le moissonnage ne puisse continuer. »

Un autre défi auquel doivent faire face les archivistes est de savoir comment fournir l'accès à ces contenus, et sous quelles conditions de droit d'auteur (les cadres juridiques variant beaucoup d'un pays à l'autre). Il est aussi essentiel de déterminer sous quel format de fichier conserver à long terme ces documents.



Réseaux sociaux © Camille Stassart

#moulefrite, le patrimoine belge à la loupe

Cette analyse a permis à l'équipe du projet de dégager certaines leçons pour mener un projet-pilote visant à archiver du contenu de Twitter et d'Instagram, présentant un intérêt pour la société belge.

Les scientifiques combinent deux approches : « La première consiste à créer des collections de données sur un thème en particulier. Nous avons déjà compilé une série de « hashtags » et de comptes Twitter et Instagram sur les commémorations belges, la gastronomie belge, les minorités en Belgique, les musées, et la communauté d'expatriés. Un des plus grands challenges est de trouver ce contenu en ligne », note Friedel Geeraert.

D'où l'intérêt de la seconde approche : s'appuyer sur les suggestions du public. Via une [campagne de crowdsourcing](#), les citoyens peuvent renseigner des comptes publics ou « hashtags » qui devraient, selon eux, être conservés dans le patrimoine belge.

« On exploite de cette manière le savoir de la communauté. Jusqu'à présent, nous avons déjà eu quelques propositions, telles que #Stoofvlees, #mannekenpis, ou encore #moulefrite. Mais aussi des comptes d'institutions culturelles, comme celui du Musée des Instrument de Musique (MIM) », énumère Fien Messens.

En parallèle, l'équipe cherche à archiver du contenu diffusé sur Twitter et Instagram relatifs aux

collections des journaux papiers de la KBR.