

# Open data et *crowdsourcing* : un état des lieux du point de vue des musées

— par Beat Estermann \*



*Backstage pass* au British Museum : des Wikipédiens sont invités à un tour guidé dans le but d'améliorer la couverture de la collection du musée sur l'encyclopédie libre.

© Mike Peel, Wikimedia Commons, CC BY-SA

Une enquête menée par la Haute école spécialisée bernoise en Suisse permet de faire le point sur la mise en place par les musées des innovations liées au développement d'Internet (numérisation des collections, réseaux sociaux, licence libre, création collaborative de contenus...) : l'auteur met en lumière la perception par les musées des avantages et des inconvénients de ces nouvelles pratiques mais aussi des obstacles rencontrés dans leur diffusion.

**D**epuis l'émergence d'Internet, des processus de changement sont en cours dans les institutions patrimoniales, ils se manifestent par des tendances successives. Ceci inclut par exemple la numérisation systématique des objets patrimoniaux, une mise en réseau accrue entre les institutions, la coproduction de contenus par les utilisateurs, l'utilisation croissante des réseaux sociaux ainsi que la mise à disposition gratuite de données et contenus sur Internet. Une étude comparative internationale analyse la position de ces institutions patrimoniales vis-à-vis de ces tendances. Le présent article présente la situation de la perspective des musées. Les données de base proviennent d'enquêtes en Pologne, en Finlande, en Suisse et aux Pays-Bas.

## Les musées en mutation

Le Derby Museum au centre de l'Angleterre mise sur les codes QR pour enrichir les explications relatives aux objets exposés à l'aide d'articles sur Wikipédia ; en plusieurs langues, bien sûr. Le musée de Brooklyn se fie au jugement des visiteurs de son site Web pour choisir des photographies pour une exposition. Le Louvre fait appel à un vaste public afin qu'il contribue à l'acquisition d'œuvres d'art. La Galerie nationale du Danemark invite à remixer une partie de son inventaire dans le cadre d'un événement "Set Art Free". Au Rijksstudio, on peut recomposer soi-même les œuvres du Rijksmuseum d'Amsterdam. Et le

\* Beat Estermann est collaborateur scientifique à l'Institut de E-Government de la Haute école spécialisée bernoise  
beat.estermann@bfh.ch

muséum de Toulouse collabore avec Wikimedia France pour photographier des objets de sa collection et mettre à disposition les photographies sur Wikimedia Commons pour permettre leur utilisation. Ce sont tous des exemples du profond changement dû à l'émergence d'Internet et aux innovations en découlant dans le domaine des musées. Concrètement, on peut identifier plusieurs tendances successives.

En Europe, depuis le début des années 2000, les institutions patrimoniales entreprennent des efforts coordonnés pour numériser de manière aussi complète que possible les objets patrimoniaux et leurs métadonnées. À côté de la mise en réseau accrue entre institutions pour l'échange de savoir-faire et éviter les doublons, des offres "Single Point of Access" émergent aussi sous forme de systèmes de catalogues inter-organisations ou de portails de musées.

Au cours de la deuxième moitié des années 2000 s'ajoute la tendance des possibilités d'interactions accrues découlant de l'émergence du Web social. Elle se manifeste à travers de nouvelles formes de personnalisation, ainsi qu'une intégration accrue des utilisateurs ou des visiteurs à des fins de coproduction. De plus, avec l'ascension de Wikipédia, Flickr et d'autres offres communautaires semblables, la création et la compilation collaboratives de contenus sur Internet se trouve de plus en plus au centre de l'attention. On assiste alors au développement d'une collaboration entre les institutions patrimoniales et les communautés en ligne ainsi qu'au lancement de projets de *crowdsourcing* indépendants par certaines institutions<sup>(1)</sup>.

À partir de 2009, les États-Unis et la Grande-Bretagne ont également lancé le mouvement *Open Government Data*, qui s'est répandu très rapidement pour inclure en peu de temps plus de 50 États dans le monde ; son objectif : les données publiques doivent être mises à disposition gratuitement sous forme lisible par ordinateur pour être utilisées par des tiers. En plus d'une vision politique institutionnelle visant la transparence, la participation et la collaboration, le mouvement *Open Government Data* est également porté par une vision technique et économique : la mise en réseau d'un maximum d'ensembles de données "ouverts" (*linked data*) doit donner naissance au "Web sémantique". Les données des autorités et le cas échéant aussi d'entreprises privées deviennent ainsi des ressources d'infrastructure d'accès libre, servant de base à des services à valeur ajoutée pour des tiers : pour la

valorisation des données, la création et la visualisation de nouveaux liens entre les données et d'autres services basés sur les données. L'utilisation de licences d'utilisation "libres" standardisées permettant la modification et la diffusion des données et contenus aussi à des fins commerciales est essentielle, car elle diminue nettement les coûts de transaction pour l'utilisation par des tiers<sup>(2)</sup>. Avec l'extension du domaine d'application de la directive européenne PSI aux bibliothèques, musées et archives décidée en 2013 – mise en œuvre dans chaque État membre de l'UE depuis l'été 2015 – le principe d'open data sera aussi ancré officiellement dans le secteur des institutions patrimoniales au sein de l'Union européenne. Étant donné que cela ne concerne pas seulement les métadonnées, mais que les institutions patrimoniales portent aussi un intérêt particulier aux licences libres pour les documents numériques relatifs aux objets patrimoniaux, on distingue dans le domaine des données culturelles les open data (par exemple les métadonnées) et l'open content (contenus numérisés).

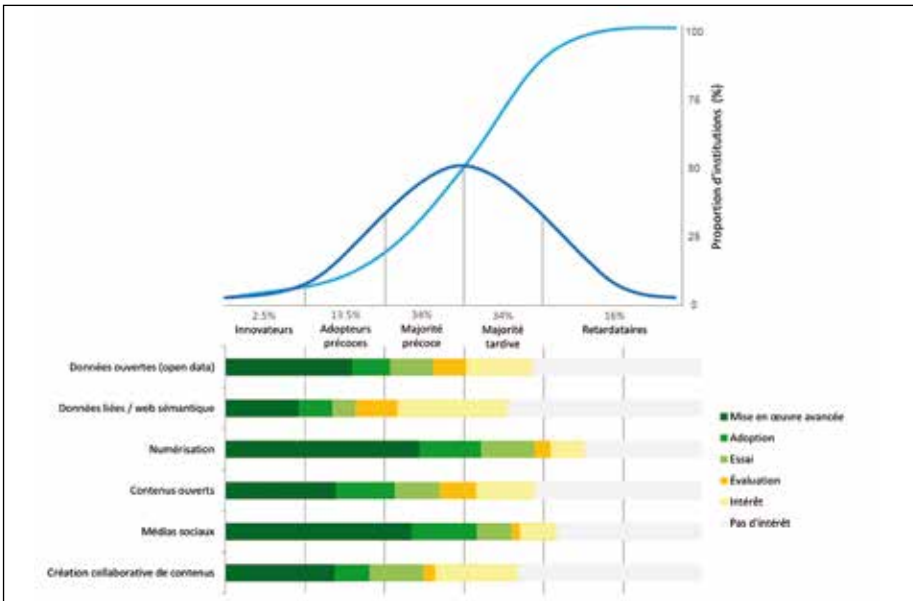
Van Go Yourself : un site web (<http://vangoyourself.com>) qui invite les utilisateurs à recréer des œuvres d'art.



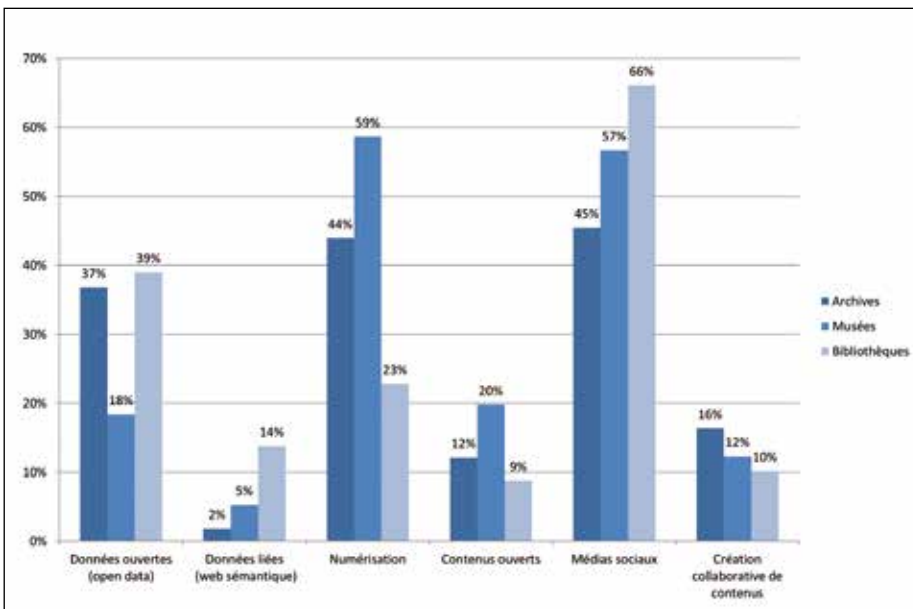
## Où se situent les musées vis-à-vis de ces développements ?

Afin de démontrer où en sont les institutions patrimoniales en matière de mise en œuvre d'open data/open content et de *crowdsourcing*, les chances et les risques qu'elles y voient et les obstacles auxquels elles font face, la Haute école spécialisée bernoise mène une enquête en ligne en coopération avec des partenaires locaux dans divers pays en 2014/2015<sup>(3)</sup>. Jusqu'à présent, on dispose des résultats de quatre pays européens. Vous trouverez ci-dessous un premier aperçu de la situation du point de vue des musées interrogés en Pologne, en Finlande, aux Pays-Bas et en Suisse<sup>(4)</sup>.

(1) Le terme « *crowdsourcing* » désigne des activités participatives en ligne pour lesquelles une personne ou une organisation demande publiquement à un groupe généralement pas précisément défini de personnes de milieux différents d'exécuter une tâche sur la base du volontariat. On parle aussi de « coproduction » ou, surtout dans le contexte de Wikipédia, d'une création collaborative de contenus. (2) Peuvent être cités comme exemples les licences de droits d'auteur Creative Commons, telles que CC-by ou CC-by-sa (pour plus d'informations, consulter : [www.creativecommons.org](http://www.creativecommons.org)) (3) Voir : <http://survey.openglam.ch> (4) Le nombre de musées ayant rempli (presque) intégralement le questionnaire est de 284 (134 en Suisse, 57 aux Pays-Bas, 51 en Finlande et 42 en Pologne). Les taux de réponse sont différents d'un pays à l'autre. Pour les musées, ils sont de 21 % en Finlande, 16 % en Suisse et 10 % en Pologne et aux Pays-Bas. Dans les quatre pays, le taux de réponse des musées est inférieur à celui des archives et bibliothèques (le taux de réponse de toutes les institutions se situe entre 12 % et 26 % selon le pays).



L'adoption des différentes pratiques par les musées, visualisée sur la base du modèle de diffusion des innovations selon Everett M. Rogers.

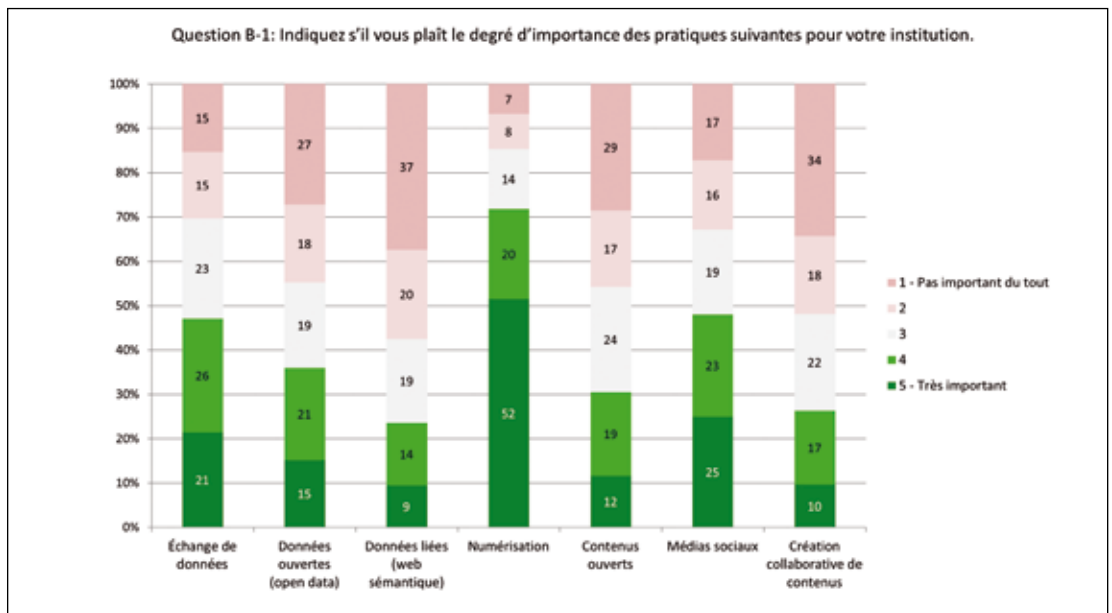


Taux d'adoption des différentes pratiques parmi les institutions patrimoniales, comparaison des différents types d'institutions.

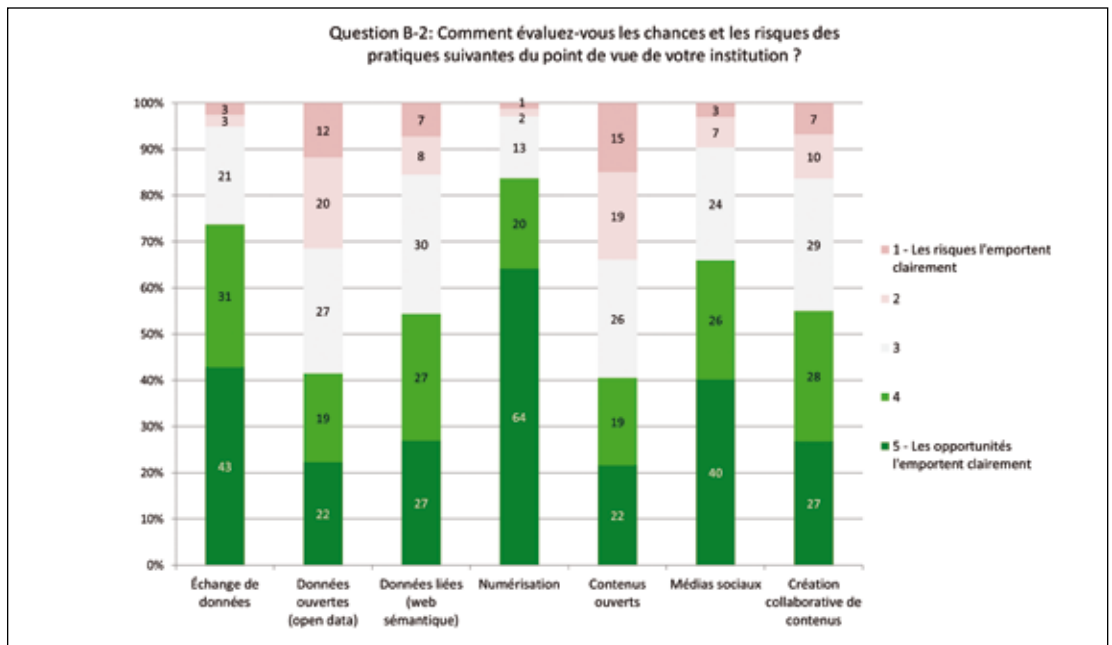
Examinons tout d'abord les taux d'adoption des pratiques innovantes étudiées : l'analyse des taux d'adoption est basée sur le modèle de diffusion des innovations popularisé par Everett M. Rogers. D'après ce modèle, la diffusion d'une innovation est un processus social dans le cadre duquel des personnes ou – comme dans le cas présent – des organisations abordent des idées ou des pratiques nouvelles pour eux et choisissent de se les approprier ou non. Chaque organisation traverse alors un processus de décision en matière d'innovation : tout d'abord, l'organisation concernée doit prendre connaissance de la nouvelle idée ou pratique et prendre position à ce sujet. Si l'innovation l'intéresse, elle va davantage s'informer à ce sujet, réfléchir aux effets concrets que cette innovation pourrait avoir sur elle-même et son environnement, et effectuer quelques essais. Si elle est satisfaite des essais, elle décidera d'introduire définitivement l'innovation (adoption). En fonction du type d'innovation, la mise en œuvre d'envergure peut durer plus ou moins longtemps. Si, dans le cadre de ce processus, l'organisation conclut que l'innovation n'est pas avantageuse pour elle, le processus de

décision en matière d'innovation est interrompu et l'innovation n'est pas mise en œuvre, ou seulement partiellement. Les premiers résultats montrent à quel point la diffusion des pratiques étudiées a déjà progressé parmi les musées : la plus répandue est la numérisation avec un taux d'adoption de 60 %, ainsi que l'utilisation des réseaux sociaux avec un taux d'adoption de 55 %. Suivent les open data et open content avec des taux d'adoption de 20 %. Ces pratiques sont en train de se répandre au-delà du cercle des *early adopters*. Par contre, avec des taux d'adoption respectifs de 10 % et 5 %, le développement collaboratif de contenus et les *linked data* n'en sont qu'à leurs débuts. Le point commun à toutes les tendances étudiées est que plus de 70 % des musées ayant répondu s'y intéressent.

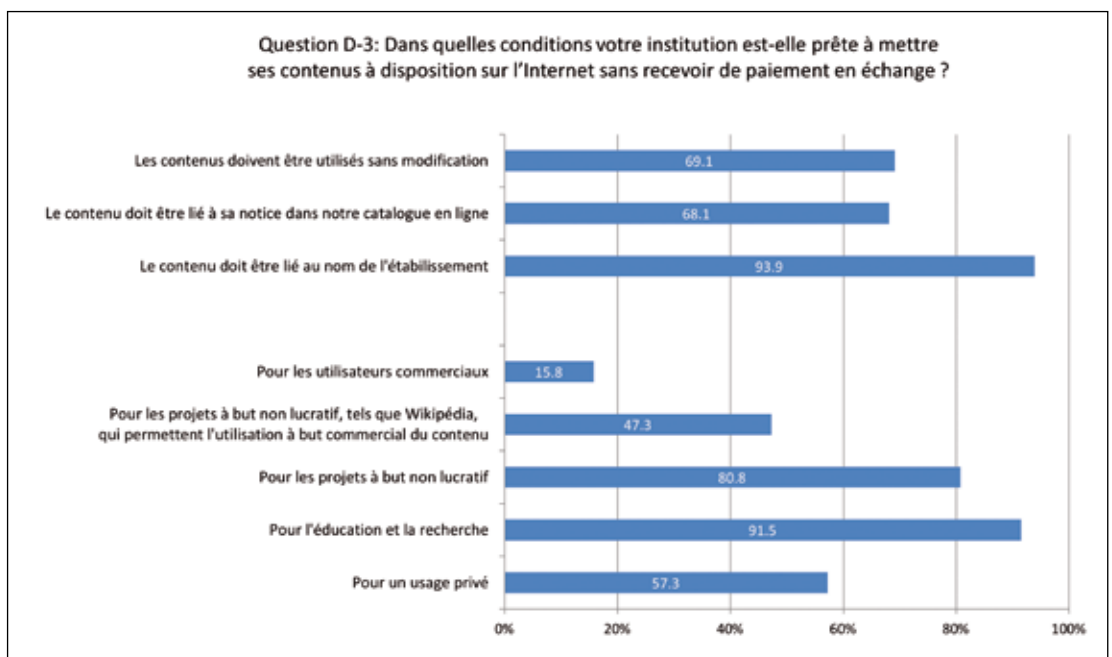
Si l'on compare les musées aux autres institutions patrimoniales, il apparaît qu'ils sont en avance en matière de numérisation et de licences libres pour les contenus par rapport aux archives et bibliothèques. Pour les open data, la situation est inversée – les bibliothèques sont en tête, alors que les musées présentent un faible taux d'adoption.



L'importance accordée aux différentes pratiques par les musées



L'évaluation des risques et opportunités des différentes pratiques du point de vue des musées



Conditions et usages selon lesquels les musées sont prêts à mettre leurs contenus numériques à disposition sur Internet.

Les bibliothèques et musées sont en tête pour l'utilisation des réseaux sociaux, alors que les archives sont un peu en retrait ; quant à la création collaborative de contenus, tous les types d'institutions présentent des taux d'adoption proches.

L'enquête permet aussi de comparer les pays entre eux. On y constate de nettes différences : en matière d'open data, les musées néerlandais sont clairement devant avec un taux d'adoption de 35 %, suivis des finlandais et polonais, qui n'ont que la moitié de cette valeur (17-18 %) puis des musées suisses, qui clôturent le classement avec un taux d'adoption de 12 %. Les *linked data* donnent une image semblable, mais l'écart entre les musées néerlandais (14 %) et les musées des autres pays (2-4 %) est proportionnellement encore plus important. Les différences sont moins nettes dans la numérisation et les licences libres, mais là encore les musées néerlandais sont en tête. Par contre, l'avance néerlandaise n'est pas visible dans l'utilisation des réseaux sociaux : la Pologne, les Pays-Bas et la Finlande sont en tête avec des taux d'adoption entre 68 % et 79 %, alors que la Suisse est nettement derrière avec 38 %. On observe le même modèle en matière de création collaborative de contenus, où la Pologne, les Pays-Bas et la Finlande présentent des taux d'adoption entre 16 % et 26 %, alors que les musées suisses occupent la dernière place avec 4 %.

D'autres facteurs d'influence ont également été analysés, comme la taille des institutions (mesurée au nombre de collaborateurs et au budget annuel), leur situation géographique et leur part de financement public : il est peu surprenant de constater que les musées locaux utilisent moins souvent les réseaux sociaux que les musées ciblés sur le niveau national ou international. De plus, les données indiquent que les grands musées ont davantage tendance à suivre une approche open data et à utiliser les réseaux sociaux que les petites institutions. Cependant, il est surprenant que les musées financés à moins de 20 % par des moyens publics aient davantage tendance à publier leurs contenus que les musées financés au moins à 80 % par des moyens publics. En effet, si l'on suit l'argumentaire des représentants du mouvement open data, les institutions à financement public devraient justement publier leurs données et contenus pour utilisation. Ceci nécessite des explications, et il sera intéressant de vérifier si ce phénomène est également présent dans les autres pays où l'enquête est encore en cours.

### La perception des différentes pratiques

Voyons à présent dans quelle mesure les musées considèrent que les pratiques analysées sont importantes et comment ils évaluent les chances et les risques associés : en matière d'évaluation de l'importance des différentes pratiques, on observe un modèle semblable aux taux

d'adoption – plus une pratique est répandue parmi les musées, plus elle paraît importante aux participants à l'enquête. L'évaluation des chances et des risques indique une image différente : on y voit les plus grands risques dans l'open data et l'open content. La perspective de l'utilisation libre des données et contenus semble donc particulièrement inquiéter les musées.

Un aperçu de la réflexion des responsables des musées répond à la question des conditions auxquelles l'institution est prête à mettre à disposition gratuitement ses contenus numérisés sur Internet (à condition qu'aucune disposition légale ne l'interdise) : les musées présentent une forte propension à mettre à disposition gratuitement des contenus à des fins de formation et de recherche, ainsi que pour des projets d'utilité publique. Mais il y a encore de nettes réserves vis-à-vis de l'utilisation privée et dès que l'utilisation commerciale entre en jeu, la disposition à publier des contenus pour utilisation chute considérablement. L'exigence largement répandue des musées interrogés que les œuvres publiées ne puissent pas être modifiées est également en contradiction éclatante avec les principes de l'open data. Paradoxalement, cette exigence, soutenue par 70 % des musées, est également largement répandue parmi les musées indiquant mettre à disposition une partie de leurs contenus sur Internet comme open content pour l'utilisation libre par des tiers. Chez les musées indiquant mettre à disposition une partie de leurs contenus comme open content, les réserves vis-à-vis de l'utilisation commerciale sont tout de même un peu moins importantes que chez les autres institutions, mais ici aussi, environ la moitié des interrogés indiquent qu'ils ne souhaitent pas mettre à disposition les contenus pour un usage commercial. Dans ce cas, on se demande si le principe de l'open content a vraiment été compris par les institutions concernées. Il semble du moins y avoir une certaine contradiction entre les idées exprimées en matière de publication sous licence libre des contenus et la pratique indiquée.

### Les chances et les risques de l'open content et du crowdsourcing

Pour plus de 80 % des musées interrogés, l'open content est un moyen important pour améliorer la perception extérieure de l'institution et pour rendre son inventaire plus accessible. D'autres motifs importants pour une approche open content sont la mise en réseau simplifiée entre institutions patrimoniales, la meilleure accessibilité des contenus pour les utilisateurs existants, l'acquisition de nouveaux utilisateurs ainsi que l'amélioration de l'échange avec les utilisateurs (70-80 % des musées interrogés partagent cet avis) et enfin, 70 % des musées interrogés indiquent que l'open content en général est un moyen important pour mieux accomplir leurs tâches principales.

Concernant les risques, 75-80 % des musées interrogés craignent une réutilisation des contenus sans attribution correcte à l'institution ou à l'auteur ainsi que des abus et de fausses présentations des contenus. La moitié des institutions ayant répondu considèrent également les éventuelles violations du droit d'auteur par des tiers ainsi que le temps nécessaire pour répondre aux demandes comme des risques importants liés à l'open content. Par contre, les éventuelles pertes de recettes ne jouent qu'un rôle minime (20 %).

En plus des risques, les institutions ont également pu s'exprimer à propos des défis de l'open content : le temps nécessaire et les coûts de la numérisation (94 % des cas) ainsi que la documentation correcte des contenus (89 %) étaient en tête. Pour la moitié des institutions, la libération des droits s'y ajoutait. De plus, les difficultés techniques (65 %) et des compétences insuffisantes du personnel (62 %) sont citées comme d'autres problèmes importants.

En matière d'usage prévu et des risques du *crowdsourcing*, seules les institutions ayant déjà des approches de *crowdsourcing* ou prévoyant d'en avoir dans l'année suivante ont été interrogées. Dans 50 % des cas, le *crowdsourcing* est utilisé pour compléter les collections et pour des tâches de contextualisation (les objets patrimoniaux sont placés dans un nouveau contexte ou des informations contextuelles sont mises à disposition). D'autres domaines d'application cités par un tiers des musées concernés sont la curation collaborative (sélection, disposition, présentation de contenus), le complément de métadonnées ainsi que le *crowdfunding*. Enfin, un quart des musées concernés cite les tâches de correction et de transcription comme autre domaine d'application du *crowdsourcing*. Pour la majorité des musées en question, il s'agit d'avoir accès à un savoir-faire extérieur (78 %), de tester de nouveaux types de relations avec les utilisateurs/visiteurs (68 %), de promouvoir la confiance et la loyauté des utilisateurs/visiteurs (62 %), de leur donner un sentiment de participation publique et de coresponsabilité (59 %) et de pouvoir effectuer certaines tâches en dépit des contraintes de ressources (53 %).



Roger Bamkin, un des inventeurs des codes QRpedia en train de placer un code QR sur un tableau explicatif au musée de Derby en Angleterre où les premiers codes QRpedia ont été mis en place en 2011. Les codes QRpedia permettent d'accéder depuis un appareil mobile à un article Wikipédia dans la langue préférée de l'utilisateur. © Roger from Derby, Wikimedia Commons, CC BY-SA

En revanche, les musées concernés voient principalement les risques et défis du *crowdsourcing* dans le fait que la préparation et le suivi sont complexes (71 %), le temps nécessaire est difficile à évaluer (65 %), la continuité à long terme de l'entretien des données n'est pas garantie (71 %) et la sécurité de la planification est limitée (57 %).

## Bilan

Les pratiques étudiées ont suscité un grand intérêt auprès des musées. Alors que plus de la moitié des musées ayant répondu numérisent déjà leurs objets patrimoniaux et utilisent les réseaux sociaux, l'open data et l'open content ne se sont jusqu'à présent imposés qu'auprès des *early adopters*. La grande majorité des musées va d'abord devoir se familiariser avec ces pratiques. Quant au *crowdsourcing* et au *linked data*, ils n'en sont qu'à leurs débuts. Il faudra probablement encore du temps avant que ces pratiques atteignent une vaste majorité des musées.

En ce qui concerne la position des musées vis-à-vis de l'open data et de l'open content, il existe encore de nettes réserves quant aux licences libres : alors que les musées ayant répondu font preuve d'une grande volonté de mettre à disposition gratuitement des contenus sur Internet à des fins de formation et de recherche ainsi que pour des projets d'utilité publique, l'idée de libérer des contenus pour l'utilisation commerciale ou d'autoriser la modification de contenus rencontre un vaste rejet. Ceci indique que dans le secteur des musées, nous sommes encore loin d'une culture de l'open data, même si les taux d'adoption de l'open data et de l'open content donnent lieu à un certain optimisme.

Alors que la plupart des musées considèrent l'open content comme un moyen important de mieux remplir leurs tâches principales, les risques et problèmes ont également un poids considérable. Le plus grand obstacle est la peur des musées face à une utilisation abusive des contenus ainsi que la crainte qu'en cas d'utilisation des contenus, l'auteur et l'institution ne soient pas dûment indiqués. De plus, le temps et les finances requis pour la numérisation et la documentation des contenus constituent un défi considérable.

Le *crowdsourcing* est perçu par les musées comme un moyen pour une participation renforcée des visiteurs et utilisateurs : alors que les musées ont l'occasion d'accéder à un savoir-faire externe et d'exécuter certaines tâches malgré le manque de moyens, il faut transmettre davantage aux visiteurs et utilisateurs le sentiment de participation et de coresponsabilité. À partir de cette nouvelle relation entre les institutions et les groupes cibles, les musées utilisant le *crowdsourcing* espèrent obtenir une confiance et une loyauté accrues de la part des visiteurs. Mais le *crowdsourcing* est généralement associé à une incertitude élevée, ce qui pourrait être l'une des raisons pour lesquelles il n'est pas encore largement utilisé. 🍷