
L'open data et les archives

Ce dossier a été coordonné par :



Cyril Longin
Archives municipales
de Saint-Étienne



Charlotte Maday
Université Paris Diderot

Introduction

L'open data, qui désigne à la fois un mouvement, une forme de réutilisation commerciale et une donnée, tend à prendre une part importante dans le discours et les actions des politiques publiques. L'État, à travers notamment la mission Étalab, mène une politique volontariste. Depuis peu, la loi NOTRe contraint les collectivités de plus de 3 500 habitants à publier leurs données existant déjà sous format électronique. La réutilisation des données publiques – et particulièrement celles conservées par les services d'archives – a beaucoup occupé, et parfois heurté, la communauté des archivistes depuis quelques années : l'« affaire » Notrefamille.com, démarrée en 2009, est encore dans tous les esprits. Pour autant, le temps d'un débat moins passionné est sans doute venu pour l'AAF. Qu'est-ce vraiment, l'open data ? Dans quels environnements juridiques évolue-t-il, notamment au regard des données à caractère personnel ? Quelle place archivistes et archives peuvent – ou doivent-ils – prendre dans cette démarche ?

Le contenu de ce dossier a fait l'objet d'une recherche et d'une réflexion préalable afin de présenter les différents points de vue de l'open data appliqué aux archives. Outre les articles proposés par un certain nombre d'organismes et de personnes qui vous sont présentés dans ce numéro, nous avons sollicité la mission Étalab, ainsi que la CNIL, pour avoir une vision « à 360 degrés » en faisant intervenir des institutions extérieures au milieu des archives, mais qui avaient, ou auraient, sur elles une influence indéniable. La mission Étalab et la CNIL n'ont cependant pas proposé d'article dans les temps impartis pour la rédaction de ce dossier.

Réutilisation et open data : quel est l'état du droit ?

La question de la « réutilisation » des archives s'est imposée à notre profession en 2009-2010 lorsqu'une société commerciale a sollicité auprès des Archives départementales la remise de dizaines de millions de fichiers reproduisant des actes d'état civil et des listes de recensement de la population pour les faire indexer nominativement à l'étranger et en revendre le résultat aux généalogistes français.

Les archivistes ont alors pris conscience que les ressources numériques produites par leurs services – les plus importantes de la sphère culturelle – pouvaient être légalement exploitées par le secteur marchand sur le fondement de la loi du 17 juillet 1978 sur l'accès aux documents administratifs, augmentée en 2005 d'un chapitre II relatif à la réutilisation des informations publiques, fruit de la transposition d'une directive européenne de 2003, alors passée inaperçue.

La loi du 17 juillet 1978 réservait pourtant aux services culturels un régime dérogatoire, mais il révéla ses limites au fil des avis de la commission d'accès aux documents administratifs (CADA) et de la jurisprudence. Signalons que celle-ci n'est pas encore consolidée, le contentieux concernant le département de la Vienne, fondé sur le droit des producteurs des bases de données, ayant fait l'objet, en juin dernier, d'un recours auprès du Conseil d'État et généré une question prioritaire de constitutionnalité.

Une nouvelle directive européenne a été adoptée le 26 juin 2013 et est en cours de transposition en droit national. Elle apporte plusieurs changements par rapport à celle de 2003 : elle intègre désormais dans son périmètre les bibliothèques, les musées et les services d'archives ; elle impose la gratuité de la réutilisation ou l'application du seul coût marginal – coût supplémentaire induit par la demande – ; elle admet néanmoins des exceptions à ce principe, notamment pour les bibliothèques, musées et archives ; enfin, elle autorise la cession de droits d'exclusivité à des tiers qui auront pris en charge des coûts de la numérisation du patrimoine culturel.

La transposition – qui devait intervenir avant le 18 juillet 2015 – utilisera un véhicule législatif spécifique, un « DDADUE » (loi portant diverses dispositions d'adaptation au droit

de l'Union européenne), et non la loi numérique en raison de son échéance plus lointaine. Le DDADUE a été présenté en Conseil des ministres le 31 juillet 2015 et devrait être examiné par le Parlement durant l'automne. Les dispositions de la directive constituent un socle minimal à respecter, les États membres pouvant aller au-delà. C'est l'option que la France a choisi afin de favoriser l'open data, mouvement dans lequel elle s'est résolument inscrite depuis 2011 avec la création d'Étalab, l'ouverture de son portail data.gouv.fr et son adhésion à l'*Open Government Partnership* (OGP), dont elle assurera la présidence en 2016-2017.

Afin de favoriser la réutilisation des informations publiques, dans un double objectif de développement économique et de transparence administrative, le projet de loi pose le principe de la gratuité de la réutilisation, alors que la directive autorisait l'application du coût marginal, et intègre dans son champ d'application tous les services et établissements culturels, d'enseignement et de recherche, quand la directive imposait seulement l'inclusion des bibliothèques, des musées et des services d'archives.

Le projet de loi prévoit néanmoins, à la demande du ministère de la Culture et de la Communication, une exception au principe de gratuité lorsque les données sont issues des programmes de numérisation des fonds et collections des bibliothèques, des musées et des services d'archives. Dans ce cas, le schéma de tarification pourra être supérieur au coût marginal et prendre en compte les coûts de collecte, de production, de conservation, de diffusion ou encore de mise à disposition. C'est une dérogation de taille, justifiée par l'effort budgétaire supplémentaire que représente, par rapport aux données nativement numériques, la numérisation du patrimoine culturel, essentiellement constitué de biens physiques non réutilisables en l'état.

Le projet de loi modifie également l'article 16 de la loi du 17 juillet 1978, en permettant aux administrations d'imposer des licences de réutilisation, « unilatérales » ou contractuelles, même en l'absence de tarification. C'est ce qu'elles faisaient déjà, mais sans fondement légal selon la CADA. Il n'en demeure pas moins que les licences ne peuvent pas apporter de restrictions au droit de réutili-

sation. La question de la légalité des licences Share-alike (partage à l'identique), adoptées par plusieurs collectivités, demeure donc posée. Au-delà de ce point précis, les licences ont une vertu pédagogique (elles rappellent le droit) et permettent de préciser les conditions de réutilisation (par exemple en imposant une forme de citation de la source).

Ce projet de loi est désormais entre les mains des parlementaires. Le nouveau dispositif devrait être adopté dans les mois à venir.

Mais ce texte ne sera peut-être pas le seul qui modifiera la loi du 17 juillet 1978. Le secrétariat d'État au numérique et Étalab envisagent en effet de réviser certaines de ces dispositions à la faveur de la loi numérique. Ces modifications pourraient toucher aussi bien son chapitre I, relatif à l'accès, que certains articles du chapitre II non modifiés par le DDADUE. Les promoteurs du projet ont notamment affirmé à plusieurs reprises vouloir « passer d'une logique de la demande à une logique de l'offre », avec obligation de publication sur Internet des documents numériques librement communicables, sur le modèle de l'amendement dit open data de la loi NOTRe, sous réserve des dispositions de la loi Informatique et Libertés en cas de présence de données à caractère personnel.

Le régime juridique de la réutilisation est donc encore en cours de construction. On peut espérer une stabilisation en 2016 avec l'adoption du DDADUE et de la loi numérique et la décision du Conseil d'État sur le dernier contentieux en cours.

Les services d'archives devront alors, au-delà des adaptations techniques de leurs règlements, licences et tarifs, proposer à leurs tutelles des choix stratégiques en matière de réutilisation. Il ne s'agit en effet, en la matière, pas seulement de questions de droit ou d'économie, mais bel et bien de la définition d'une politique culturelle de long terme.



Bruno Ricard

Sous-directeur de la communication et de la valorisation des archives
Service interministériel des Archives de France

Le regard d'OpenDataSoft



Créée en 2011, OpenDataSoft développe une plateforme en ligne complète conçue pour la transformation rapide et économique de tout type de données en services innovants (APIs, visualisations de données ou monitoring en temps réel) sans aucune limite de volumétrie. Sa mission : faciliter la publication, le partage et la réutilisation de données par les utilisateurs métiers.

Basée à Paris, OpenDataSoft compte parmi ses clients Total, AXA, le ministère de l'Enseignement supérieur et de la Recherche, la Région Île-de-France, la Ville de Bruxelles, SNCF, la RATP, Veolia ou encore m2ocity¹.

David Thoumas, co-fondateur et directeur technique d'OpenDataSoft, répond à nos questions.

Quelle connaissance avez-vous de la notion de *records management* ?

La gestion des documents d'archives est associée pour moi à la notion de connaissance et de conservation, fiable et sécurisée, de données structurées ou non structurées, à valeur juridique, générées par une administration ou une entreprise.

Quelle connaissance et quelle vision avez-vous du monde des archives et des archivistes ?

Évoluant depuis plusieurs années dans le domaine de l'open data (OpenDataSoft développe et opère une plateforme de données qui peut, entre autre, être utilisée par les administrations et les collectivités pour exposer des données publiques), je n'ai paradoxalement que peu ou pas été en contact avec le monde des archives. Ce qui, après réflexion, me semble surprenant. J'ai l'impression donc (et évidemment ce n'est que mon point de vue) que les archivistes ne sont pas encore tout à fait partie prenante des projets open data.

Le mouvement open data nécessite des données de qualité (structurées, qualifiées, authentiques et fiables). Quelle place donnez-vous aux archives comme information ? Quelle place, selon vous, peuvent ou doivent occuper les archivistes au regard de ce mouvement ?

Il est indispensable que, progressivement, les plans d'action open data incluent les archivistes et le monde des archives en général. Les projets open data se sont jusqu'à présent

concentrés sur la mise à disposition d'une certaine typologie de données :

- données géographiques (extension logique des politiques SIG) et pour lesquels le public a un fort intérêt ;
- données de budget ;
- données politiques (élections, etc.).

Il y a aujourd'hui clairement un manque dans la diffusion des données administratives/juridiques officielles (données courantes et historiques). Les archivistes sont probablement les mieux placés pour traiter ce sujet.

Par ailleurs, les projets open data sont souvent portés par une entité particulière : l'équipe SIG, la communication, la direction des systèmes d'information (DSI). Les archivistes apporteraient sans aucun doute une dimension transversale aux équipes de gouvernance des projets open data.

Enfin, la plateforme open data est un média d'échange avec le grand public, qui permet notamment de mettre en contact les réutilisateurs avec les producteurs. Il y a donc un intérêt certain pour les archivistes à se positionner au cœur de la démarche, pour la valorisation de leurs métiers mais également une meilleure compréhension des besoins du grand public.

Les archivistes produisent des données qualifiées lors de la collecte et du classement (identification du producteur, contexte de production, communicabilité, etc.). Parallèlement, un système d'archivage électronique (SAE) garantit l'authenticité, l'intégrité et donc la fiabilité des données. Dans quelle mesure pensez-vous que nos systèmes d'information archivistiques et qu'un SAE puissent alimenter une plateforme open data ?

Une des particularités de la plateforme OpenDataSoft est de proposer la possibilité d'une connexion directe aux systèmes d'informations pour l'alimentation des jeux de données. Il me semble donc tout à fait intéressant de considérer les SAE comme source potentielle de données pour une plateforme open data.

Quels conseils donneriez-vous aux archivistes qui souhaiteraient s'investir dans ces projets ?

N'hésitez pas. L'open data est un excellent moyen pour valoriser un métier. Au-delà de l'aspect transparence, faire de l'open data est également un réel acte de communication. Par ailleurs, il est maintenant globalement établi que les principaux réutilisateurs de l'open data sont les collectivités elles-mêmes. Plus l'information est qualifiée et de qualité, plus l'intérêt pour la donnée publiée sur la plateforme open data se renforce et incidemment donne de la visibilité aux producteurs.

Pensez-vous que les données produites par les services d'archives (fiches de description des archives, documents numérisés comme l'état civil, plans, etc.) puissent alimenter les plateformes open data ? Si oui, en quoi peuvent-elles intéresser des développeurs et entreprises ?

Je suis persuadé que les documents produits par les services d'archives ont un intérêt. Je pars du principe que toute documentation ouverte au grand public peut avoir un intérêt. Il faut néanmoins bien choisir ses combats et ne pas chercher à ouvrir tout azimut. Il faut avant tout chercher à répondre aux besoins exprimés par les citoyens/ré-utilisateurs.

En tant que réutilisateur de données, je serais notamment tout à fait intéressé par tout ce qui peut être attaché à l'histoire de la cité (évolutions cadastrales, plans, etc.). Également par toutes les archives juridiques/administratives. Avec un challenge supplémentaire pour ces dernières : leur structuration.

1. Vous trouverez plus d'informations sur www.opendatasoft.com ou sur notre compte Twitter : @opendatasoft.

Le *data management plan* : un outil pour l'*open access* et l'archivage des données de recherche

En décembre 2013, la Commission européenne lançait un projet pilote sur l'ouverture des données de recherche dans le cadre de son programme de financement Horizon 2020. Les chercheurs qui postulent à certains appels à projets doivent maintenant diffuser les données de résultats en appui de leurs publications et s'assurer de leur conservation pérenne.

Ce pilote s'inscrit dans l'engagement de la Commission européenne en faveur de l'*open access* ou libre accès, un mouvement né dans les années 1990 pour faciliter les échanges scientifiques et qui se concrétise en 2003 par la « Déclaration de Berlin ». Son objectif est d'offrir « de nouvelles possibilités de diffusion de la connaissance¹ », en dehors du circuit classique des revues scientifiques payantes. Par ses principes de libre échange et de transparence, l'*open access* préfigure l'*open data* qui devient son pendant pour les données administratives.

Consciente que seules les données documentées peuvent être réutilisées, la Commission européenne demande aux chercheurs de se doter d'un *data management plan*, ou plan de gestion de données, qui les accompagnera au cours de leur recherche. Il s'agit d'un outil qui décrit les données collectées et produites, les modalités de stockage au cours du projet, ainsi que les conditions de diffusion et de conservation à l'issue du projet. Il est conçu comme un document évolutif qui invite les chercheurs à réfléchir à la gestion de leurs données et constitue un livrable rendu à la Commission européenne lors de l'évaluation finale.

Le lancement du pilote augure de grands changements dans la collecte des archives de la recherche. Jusqu'ici, celle-ci intervenait tard dans la carrière du chercheur et l'archiviste collectait souvent des documents que le chercheur n'avait pas encore détruits ou légués à un collègue. Avec le pilote de la Commission, un organisme financeur demande pour la première fois aux chercheurs de s'intéresser à la conservation pérenne de leurs données et leur suggère un outil pour le faire : le plan de gestion de données. L'archivage peut donc être anticipé dès le lance-

ment du projet de recherche, ouvrant la voie à une forme de *records management* du processus de recherche.

Afin d'accompagner les chercheurs, les universités Paris Descartes et Paris Diderot ont constitué un groupe de travail composé d'un conservateur des bibliothèques de l'université Paris Descartes, d'une documentaliste et d'une archiviste de l'université Paris Diderot. Ce groupe de travail a conduit à la réalisation d'un guide de rédaction de plans de gestion de données pluridisciplinaires, structurés de la façon suivante :

- **Section 1.** Informations sur le projet
- **Section 2.** Responsabilités des données (nom du responsable, propriété des données)
- **Section 3.** Ressources nécessaires à la mise en œuvre du DMP
- **Section 4.** Jeux de données
 - Section 4. 1. Documentation des données
 - Section 4. 2. Au cours du projet : stockage, accès et sécurité des données
 - Section 4. 3. Métadonnées : documentation et organisation des données
 - Section 4. 4. À l'issue du projet : dissémination des jeux de données
- **Section 5 :** Sélection et archivage [à l'échelle du projet]²

Les problématiques de diffusion et d'archivage des données de recherche y sont abordées conjointement, mais ne se pratiquent pas au même niveau. La diffusion intervient au niveau du jeu de données, que nous avons défini dans ce guide comme « un ensemble de données techniquement homogène ou intellectuellement cohérent identifié comme tel ». Ce jeu de données est souvent un en-

semble de résultats qui vient en appui d'une publication. L'archivage, quant à lui, se réalise sur la base de l'ensemble des données produites au cours du projet selon la logique du respect des fonds.

La diffusion comme l'archivage répondent à la même logique – donner accès – dans des temporalités différentes. La clé de la réussite des plans de gestion de données est donc la coopération entre les professions : les documentalistes qui accompagnent les chercheurs au quotidien dans la description de leur production scientifique, les bibliothécaires spécialistes de la diffusion documentaire, et les archivistes qui apportent leurs compétences en gestion du cycle de vie et en conservation. Les métiers de l'information scientifique et technique doivent donc œuvrer de concert au service du producteur – l'équipe de recherche – afin de valoriser son travail.



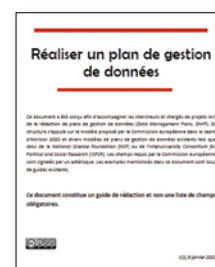
Aurore Cartier
Conservateur des bibliothèques
Université Paris Descartes



Magalie Moysan
Archiviste
Université Paris Diderot



Nathalie Reymonet
Documentaliste à la direction
de la recherche
Université Paris Diderot



1. Déclaration de Berlin, traduit par l'Institut de l'information scientifique et technique (INIST) du CNRS, accessible à cette adresse : <http://openaccess.inist.fr/?Declaration-de-Berlin-sur-le-Libre>

2. Accessible sous licence CC-BY-NC-SA à cette adresse : <http://www.sorbonne-paris-cite.fr/index.php/fr/recherche/reseau-recherche-europe/778-realiser-un-plan-de-gestion-de-donnees-de-la-recherche-data-management-plan-dmp-guide-de-redaction>

Archives et open data : quels points de vue ?



Pascale Verdier, directrice des Archives départementales du Bas-Rhin, répond à nos questions.

Quel regard portez-vous sur le mouvement open data en général ?

En tant que fonctionnaire, je ne suis pas certaine d'avoir à porter quelque regard que ce soit sur le sujet – en tant que citoyenne c'est naturellement autre chose.

L'open data caractérise, selon Wikipédia, « une donnée numérique [...] diffusée de manière structurée selon une méthode et une licence ouverte garantissant son libre accès et sa réutilisation par tous, sans restriction technique, juridique ou financière. L'ouverture des données (en anglais *open data*) représente à la fois un mouvement, une philosophie d'accès à l'information et une pratique de publication de données librement accessibles et exploitables ».

Outre les principes de transparence et de démocratie, souvent avancés, je constate qu'il s'agit surtout, en fait, pour des entreprises privées, d'obtenir des données gratuitement, puis de les revendre (alors même qu'elles revendiquent le fait de ne pas dépendre de l'État providence).

En même temps qu'open data, me vient à l'esprit le mot « enclosure » qui était en usage au Moyen Âge quand ont été en quelque sorte « privatisés » avant l'heure certains biens, dits communs. On a en effet reproché aux services d'archives n'ayant pas mis en place des licences dites libres de fermer des biens communs, de revenir aux enclosures. Or, pour moi, le « bien commun » c'est le document original, pas forcément sa copie numérique. « Commun » vient d'ailleurs du latin « *munus* », qui signifie à la fois « don » et « charge ». Autrement dit, recevoir un *munus*, c'est aussi se trouver « obligé » à quelque chose en échange.

L'open data, sous ce prisme, n'aurait vraiment de sens que si les entreprises, à leur tour, diffusaient gratuitement des données obtenues gratuitement. Quelle réciprocité, quel « *munus* » va-t-on consentir aux établissements

qui engagent des frais conséquents pour numériser des documents qui pourraient n'être accessibles qu'en salle de lecture ?

Pour moi, outre ces positions de principe, la principale question est la suivante : toutes les données ont-elles vocation à être « librement accessibles et exploitables » par tout un chacun ? À mon avis, la réponse est clairement non, au moins lorsqu'on parle de données personnelles. Avons-nous vraiment envie de voir développées des bases de données nominatives géantes, qui permettraient de savoir combien on compte de repris de justice ou de maladies génétiques dans une même famille ?

Au regard de ce mouvement, quelle place accordez-vous aux établissements culturels ?

Les établissements culturels ont actuellement une place à part, prévue par la loi. Ce n'est pas moi qui la leur accorde.

Il se trouve qu'en matière de numérisation, les archives occupent une part extrêmement importante dans la production de données numériques – bien plus que beaucoup d'autres institutions culturelles. Or la numérisation est le fruit d'une politique volontariste, et l'on pourrait tout à fait imaginer qu'elle cesse, par exemple si les difficultés budgétaires des collectivités continuaient de croître.

Les données numérisées jusqu'à présent, qui comportent essentiellement des données personnelles, sont considérées comme « rentables » par le secteur privé, qui souhaite les obtenir le moins cher possible. C'est pour cela que les Archives ont été, en France, en quelque sorte « pionnières » dans la mise en place des licences.

Pourquoi les données produites par les personnes publiques devraient-elles être nécessairement mises gratuitement à disposition d'un usage commercial, qui vise à les « enfermer » dans un système payant ? On est à nouveau bien loin du « bien commun ». On me dit : parce qu'on les a payées avec les impôts.

Ce sont aussi nos impôts qui financent les piscines municipales, ou l'eau, et y avons-nous accès gratuitement ?

On numérise pour plusieurs raisons : parce que le document est très consulté (registres paroissiaux ou états signalétiques et des services), parce qu'il est très fragile (plaques de verre ou fonds anciens), difficile à consulter en salle de lecture (plans cadastraux) ou présente des risques élevés de vol. Dans ces hypothèses, là où auparavant on microfilmaient, on numérise à présent, les technologies ayant évolué. Le rôle du service public n'est pas de suivre seulement la demande du public le plus nombreux, il est aussi de protéger les fonds risquant la détérioration. Nous allons donc probablement, dans les prochains temps, numériser d'autres fonds que ceux comportant des données personnelles.

Ceci pour dire que toutes les données, loin s'en faut, ne sont pas nativement numériques : entre des horaires de bus, que quelqu'un collecte de toute manière sous forme de tableur, les « zones accidentogènes » (également disponibles nativement sous forme de base de données) et les données qu'il a fallu ordonner et sélectionner, avant de les numériser, il y a une grande différence, et des coûts extrêmement variables également.

Pour terminer sur ce point des établissements culturels, j'observe que si la BnF met en ligne gratuitement certaines données, dans les bibliothèques il faut, le plus souvent, pour emprunter, s'acquitter d'un droit d'accès. Les musées font également payer des entrées, et les produits en ligne sont également vendus (applications pour une exposition par exemple) : les développements gratuits sont souvent indigents. Quant à l'Inventaire du patrimoine, il obtient gratuitement des images qui sont ensuite intégrées à ses publications payantes. Seuls les services d'archives sont totalement gratuits (je crois qu'il n'y a plus qu'un seul site payant en France, celui du Calvados).

S'agissant de la réutilisation des archives, le Bas-Rhin a adopté un règlement, des licences ainsi qu'une redevance en cas de diffusion publique commerciale. Quels ont été les motifs de cette décision ?

Les motifs immédiats ont été de mettre en place ce que la loi permettait, c'est-à-dire des tarifs en cas de réutilisation commerciale. Je rappelle que tous les documents actuellement en ligne, soit bien plus d'un million d'images, sont accessibles gratuitement sur le site des Archives départementales. Si on veut les diffuser publiquement et à titre commercial (il faut que les deux critères soient cumulés), il y a redevance de réutilisation. Encore y a-t-il des exonérations en fonction du tirage des publications, pour les établissements culturels publics, etc.

Pourquoi ? Il ne me semble pas particulièrement « moral » que des entreprises privées mettent la main gratuitement sur des données dont la réalisation a généré des coûts très importants – plus d'un million d'euros dans le Bas-Rhin –, même si elles mettent en avant les créations d'emplois (surtout à l'étranger) générées par l'indexation de ces actes – alors même que certaines entreprises avaient également acheté les index réalisés par les bénévoles d'associations généalogiques. Je me demande donc de quelle indexation il s'agit exactement, et je continue à ne pas voir clairement quelles sont les richesses ainsi créées dans notre pays.

Qu'offrent ces entreprises que les services d'archives n'offrent pas ? Une indexation nominative. Pourquoi, puisqu'il s'agit effectivement d'une demande sociétale, les Archives n'y répondent-elles pas ? Parce que ce n'est pas leur mission, d'abord, et ensuite parce qu'elles ont une nécessaire exemplarité scientifique à respecter, ce qui n'est pas le cas de toutes ces entreprises.

En quoi cela me dérange-t-il qu'une entreprise vende ce que mon site offre gratuitement ? Il y a un risque réel de chute de fréquentation des sites Internet publics : celui du Bas-Rhin est devenu le plus fréquenté, ou peu s'en faut, de la région. Ce qui a donné aux Archives une vraie légitimité aux yeux des élus, légitimité qu'il perdrait en cas de baisse drastique de fréquentation.

Quelles ont été, ou sont encore, les conséquences de ce choix ?

La seule réutilisation payante est celle qui consiste à diffuser publiquement des images dans un but lucratif, les établissements culturels publics étant exemptés, de même que les publications à faible tirage. Je ne pense pas que la majorité des chercheurs, des utilisateurs d'images, soient pénalisés par ce choix.

Avez-vous observé une hausse ou une baisse des sollicitations ?

Une hausse, et j'en suis très contente, car je considère que par ce biais les gens comprennent que ces images ne leur appartiennent pas, tout en pouvant les utiliser.

Ne pensez-vous pas que soumettre la réutilisation commerciale à une redevance puisse limiter la réutilisation ?

Poser la question de cette manière, c'est y répondre ! Plus sérieusement, le règlement et les licences ne soumettent pas toute réutilisation commerciale à redevance – seulement la diffusion publique d'images à but commercial, c'est bien plus restreint.

On pourrait poser d'autres questions : pourquoi ces entreprises veulent-elles à toute force détenir les images ? Pourquoi un simple lien vers le site des Archives départementales ne leur convient-il pas ? Ce serait gratuit pour le réutilisateur, générerait du trafic sur mon site, et la question serait résolue pour tout le monde. Je n'ai jamais eu la réponse à cette question, que je pose depuis le début.

Au regard du contexte réglementaire qui pourrait bouger (transposition de la directive européenne sur les données publiques) mais également du mouvement open data, votre collectivité envisage-t-elle de modifier son choix ?

Ma collectivité se déterminera naturellement en fonction de la loi. La position adoptée ici est légale, même si elle ne plaît pas. Si elle n'était pas légale, elle aurait été attaquée : je me borne à constater que cela n'a pas été le cas. Donc si la loi change, et si la position adoptée jusqu'à présent n'est plus légale, je ne doute pas que la position de ma collectivité changera. Nous ne sommes pas hors la loi.

Quel regard portez-vous sur les choix de certains services d'archives comme les Archives départementales des Yvelines et du Rhône (libre réutilisation et usage de la licence libre d'Étalab), les Archives départementales de la Vendée mais également sur celui des Archives de Toulouse (principe du partage à l'identique) ?

J'observe que les Yvelines ont exclu les données personnelles de leur licence dite « libre ». Le dispositif vendéen me semble bien compliqué à mettre en œuvre. Quant à Toulouse, je considère que ce principe de « partage à l'identique » interdit, de fait – tout en proclamant hautement le contraire – la réutilisation commerciale, qui est une volonté du législateur. Je vous mets au défi de trouver un seul usage de réutilisation commerciale dans les archives rendu possible avec ce type de licence. Pour moi, cette posture n'est juridiquement pas tenable : il ne suffit pas de proclamer un principe, il faut que celui-ci soit suivi d'une mise en œuvre. Si cette licence était appliquée partout en l'état, l'Inventaire du patrimoine ne pourrait pas facturer, quand il obtient les images d'archives gratuitement, les publications dans lesquelles elles sont publiées – c'est pourtant cette licence que certains services de l'Inventaire ont choisie pour leurs propres données : il y a là une logique qui m'échappe.

Pensez-vous que l'AAF, en tant qu'association professionnelle, ait à nourrir le débat et éventuellement faire des recommandations ?

Des recommandations, non. De plus, chacun s'est plus ou moins organisé, à une époque où les services d'archives étaient bien seuls. Les collectivités s'administrent d'ailleurs librement. Nourrir le débat, bien sûr !

Aux Archives de Toulouse on partage à l'identique

Les Archives de Toulouse se présentent comme un service administratif à vocation culturelle, fier de la richesse de ses fonds d'environ treize kilomètres, à la fois réglementaires et patrimoniaux, engagé dans la transmission et la valorisation de cet héritage.

Depuis l'automne 2013, les Archives ont choisi d'ouvrir l'accès à leurs données, après un passage de deux ans par un accès réglementé avec redevances. Cette décision mûrement réfléchie s'appuyait sur une résolution pragmatique d'alléger les tâches administratives liées à la gestion de la mise à disposition de nos ressources, mais également sur une volonté d'ouverture de l'accès à ce patrimoine commun que constituent les archives de la ville.

Avec l'aide de certains collègues, dont Jordi Navarro, et des affaires juridiques de la collectivité, au terme d'une lente maturation d'environ neuf mois, le nouveau règlement et les trois licences correspondantes ont été validés par le conseil municipal du 20 septembre 2013, permettant ainsi aux Archives d'entrer pleinement dans la politique d'open data menée par la collectivité depuis 2010. Ce vote représentait, en outre, une mise en conformité de nos actions puisque les Archives de la ville avaient été les premières, avec le Muséum d'histoire naturelle, à se lancer dans la collaboration avec Wikimedia France suite à la convention signée par la ville de Toulouse en 2010.

Le règlement définit trois types de données pouvant être qualifiés schématiquement de la manière suivante : les informations publiques dans le respect des données personnelles, les œuvres relevant du code de la propriété intellectuelle et les bases de données. Le principe pour chacune des catégories est le même : liberté de partager, de créer et d'adapter, tout en mentionnant la paternité, en partageant aux conditions identiques et en gardant ouvert. Chaque licence tacite est liée aux catégories précédemment définies. Sont bien évidemment exclues les données dont la Ville n'est pas le seul titulaire des droits. C'est la raison pour laquelle les données issues d'Urban-Hist, le patrimoine à la carte de Toulouse, ne sont pas intégralement libérées. Ce programme étant le fruit d'un partenariat entre la Ville, la métropole, la Région Midi-Pyrénées et l'État, les données sont progressivement ouvertes dès que les partenaires l'autorisent. En 2014, un jeu d'environ 3 000 fiches a été transféré sur la plateforme open data de la collectivité, la Région Midi-Pyrénées ayant accepté que les données Toulouse de l'inventaire architectural et mobilier de la Ville entrent dans le processus.



Muséum d'histoire naturelle, atelier de taxidermie. Vers 1898, Eugène Trutat.. Vue d'ensemble de l'atelier où officie Victor Bonhenry. Au premier plan, table en bois, panier de paille, animaux empaillés ; au second plan, de gauche à droite, un homme assis vue de face vêtu d'un tablier, homme debout plan américain portant costume et casquette et Bonhenry debout de face empaillant un singe sur la table ; en arrière-plan suspendus au plafond et au mur animaux empaillés ; cage ; chaises ; outils. © Ville de Toulouse, Archives municipales, dépôt de l'association « Les Toulousains de Toulouse et Amis du Vieux Toulouse », 51 Fi 206.

Depuis que la réutilisation d'une grande partie de nos fonds est permise, différentes expérimentations ont été accomplies :

- près d'un téraoctet d'images et métadonnées fournies à FamilySearch (registres paroissiaux, état civil ou recensements), certaines données remises à l'Entraide généalogique du Midi toulousain (EGMT) ;
- des jeux de données versées sur la plateforme de la collectivité, dont une liste des délibérations de 1909 à 1969 ;
- 2 614 images accompagnées de leurs métadonnées intégrées sur Wikimedia Commons (fonds Trutat sur la ville fin XIX^e-début XX^e siècles et fonds Berthelé sur l'arrière-front de 1915 à 1918) ;
- des albums, pour la plupart à destination des enseignants, sur Flickr ;
- des milliers de fichiers fournis à la demande. Les fichiers des documents numérisés sont fournis gratuitement mais le coût de numérisation à la demande est de deux euros.

L'organisation administrative s'est allégée mais la question de la gestion des droits d'auteur reste complexe en raison du manque d'information relative aux auteurs et aux droits afférents pour certaines œuvres conservées. Parfois, il est également difficile de déterminer si un document doit être considéré comme une œuvre ou non. Pour l'avenir nous réfléchissons à encore simplifier le cadre réglementaire, à intensifier nos échanges avec FamilySearch ou d'autres pour développer l'indexation collaborative notamment, et à diffuser d'autres jeux de données sur la plateforme de la collectivité...



Catherine Bernard

Adjointe du directeur des Archives municipales de Toulouse
Responsable du service des publics, fonds clos et audiovisuels

— Conclusion

L'environnement numérique bouleverse les pratiques des archivistes : avec le contexte de l'open data, on a notamment pu découvrir une nouvelle dimension à la valeur économique d'une archive, d'un document ou d'une donnée à travers la réutilisation. Bien que le droit soit encore mouvant, notamment le droit autour de la donnée, et que les pratiques dans les archives varient encore beaucoup (en fonction des contextes locaux ou des aspirations des archivistes), la réutilisation, dont l'open data, constitue une autre évolution importante, qui nous semble devoir être maîtrisée par l'ensemble de la profession. Ne pourrait-on d'ailleurs pas plutôt parler de révolution ? Jusqu'à présent, on conserve pour communiquer, dorénavant, nos tableaux de gestion – ou plans de gestion de données – devront prendre en compte les conditions et modalités de réutilisation, qu'elles soient libres, soumises à licence ou redevance. Les archivistes ont toute leur place à prendre dans la gouvernance des données, en faisant reconnaître leur expertise auprès des acteurs de l'open data et des décideurs.