



Quand des établissements de conservation du patrimoine mobilier se retrouvent sur www.e-corpus.org

MONSEIGNEUR PAUL CANART

Biblioteca Apostolica Vaticana

AVEC LA COLLABORATION DE

CAROL GIORDANO

Centre de conservation du livre

Paul Canart est vice-préfet honoraire de la Biblioteca Apostolica Vaticana (Cité du Vatican), où il a dirigé le département des manuscrits. Auteur de nombreux livres et articles, il est membre de l'Accademia dei Lincei, du Comité international de paléographie grecque, du comité de direction de la revue Scriptorium et du comité scientifique du Centre de conservation du livre.

Carol Giordano est chargé de mission au Centre de conservation du livre sur la bibliothèque numérique e-corpus.

Même si les traditions professionnelles des archives, des bibliothèques et des musées diffèrent largement, ces trois types d'établissements n'en conservent pas moins des éléments patrimoniaux peu éloignés les uns des autres.

De surcroît, il n'est pas rare de voir des bibliothèques conserver des archives sur leurs étagères, les archives conserver des tableaux et divers objets d'art et la plupart des musées conserver, en dehors de leurs propres bibliothèques et de leurs propres archives, des collections d'estampes et de livres. Il n'est pas non plus chose extraordinaire de consulter au Metropolitan Museum des imprimés du xvi^e siècle, ou de contempler des tableaux dans des bibliothèques anciennes d'Italie.

Ces objets atypiques ne sont pas sans poser des problèmes aux responsables des collections et notamment aux bibliothécaires. Faut-il, par exemple, les traiter comme des livres ou différemment ? Au-delà de cet aspect technique, et ayant moi-même effectué toute ma carrière à la Bibliothèque vaticane, je sais à quel point la coopération entre les bibliothèques, les archives et les musées représente un enjeu majeur mais aussi une difficulté quotidienne. Les objets conservés, les missions, les publics sont si différents qu'au fil de l'histoire ces institutions se sont construites de manière indépendante et dissociée pour aboutir à trois structures distinctes, tout au moins en Occident. Pour le sujet qui nous intéresse, la diffusion

électronique, il faut bien constater que les projets communs sont rares. Dans le cas du Vatican, il existe trois systèmes d'information différents remplissant chacun une fonction propre. Cela dit, il est intéressant de s'interroger a posteriori et de réfléchir aux enjeux d'une éventuelle coopération.

Le cas du Vatican mis à part, il est clair que tous les musées, archives et bibliothèques ne sont pas dotés en collections, personnels et moyens financiers comme le sont certaines grandes institutions nationales. Nombre d'institutions régionales et locales cherchent aussi à atteindre de nouveaux publics et une meilleure visibilité à travers ce qu'il est convenu d'appeler désormais « la toile » et peu de ces projets sont menés en coopération entre archives, bibliothèques et musées¹.

L'exemple que j'aimerais présenter ici est celui d'une plateforme collective, e-corpus², qui présente, à mon sens, des avantages innovants, originaux et particulièrement intéressants pour la question spécifique de la collaboration patrimoniale. En effet, si les techniques et les normes d'inventaire et de catalogage sont très distinctes d'un type d'établissement à l'autre,

1. Mentionnons néanmoins quelques projets intéressants : BAMPortal (Allemagne), Norwegian Digital Library Initiative (Norvège), Museum and Online Archives of California (USA), etc., mais reconnaissons que ces projets sont loin d'être majoritaires.

2. www.e-corpus.org



Page d'accueil de la plateforme collective e-corpus

elles se sont fortement rapprochées avec l'avènement de l'ère numérique et le développement de l'informatique (citons l'exemple de l'introduction de l'ISAD/G³ pour le catalogage des manuscrits en bibliothèque). Envisager des bases de données et d'images communes, véritables points d'accès aux références – notamment dans le domaine scientifique –, devient non seulement plus aisé pour ces établissements, mais aussi souvent le seul moyen de réussir à mettre en œuvre des projets. C'est aussi, et surtout, le meilleur moyen de répondre aux attentes des usagers potentiels de ces ressources. Hormis quelques grosses structures de diffusion, comme par exemple Europeana⁴, les projets de coopération et de création de collections numériques communes et/ou partagées entre ces trois acteurs culturels clés sont encore peu nombreux.

Présentation d'e-corpus

Prolongement d'Internum, projet opérationnel depuis plus de dix ans, e-corpus est une plateforme numérique collective et patrimoniale lancée en 2010 par le Centre de conservation du livre d'Arles⁵. Elle signale, répertorie, diffuse et permet l'accès à plusieurs dizaines de milliers d'objets et documents appartenant principalement au patrimoine écrit et linguistique (manuscrits, archives, livres, journaux), ainsi qu'au patrimoine iconographique et à d'autres objets culturels (photographies, estampes, enregistrements sonores, vidéos, objets d'art).

Gérée par une structure qui apporte sa compétence en matière de pilotage et de collaboration, elle associe depuis ses débuts les établissements qui conservent les collections patrimoniales et documentaires : bibliothèques, archives, musées et laboratoires de recherche participent au projet depuis la première étude de faisabilité, à partir de 1999. Cette première phase, Internum, était un projet européen de grande envergure dont

le but était de constituer une bibliothèque en ligne d'ouvrages numérisés de l'Europe méditerranéenne, afin que le plus grand nombre puisse y avoir accès gratuitement, grâce au réseau internet et aux nouvelles technologies de l'information et de la communication.

Bibliothécaires, archivistes, documentalistes, chercheurs et conservateurs de musées de six pays méditerranéens, accompagnés d'un comité d'experts des établissements de la Région Provence-Alpes-Côte d'Azur, ont activement travaillé et réfléchi à la mise en commun de leurs données descriptives d'objets différents. Ce défi a été relevé, et la plateforme numérique a été ouverte sur internet en 2000.

À côté d'Internum ont été par ailleurs ouvertes, à partir de la même base, deux plateformes collaboratives thématiques menées à bien dans le cadre de deux projets européens distincts, Bivimed (Bibliothèque virtuelle de la Méditerranée) et Eurindia, aujourd'hui réunies dans e-corpus. L'Unesco ne s'y est pas trompé, et a su soutenir cette dynamique en finançant également cette phase de faisabilité dans le cadre du projet Medlib. Il convient de rappeler qu'à l'époque

3. ISAD(G) : Norme générale et internationale de description archivistique; en ligne : www.icacds.org.uk/fr/ISAD%2&G%29.pdf

4. www.europeana.eu/portal

5. www.ccl-fr.org

l'enjeu du multilinguisme et du multi-alphabétisme n'était pas simple à résoudre, et les polices Unicode⁶ sont heureusement venues au secours de ce projet ambitieux.

Cependant, ces sites internet, en tant que vecteurs de diffusion, montrèrent peu à peu leurs limites. En effet, pour qu'une telle richesse documentaire soit mise à disposition des usagers, il était nécessaire d'ouvrir le système et de créer plusieurs modes d'accès : par des sites spécialisés, des collections thématiques et un travail collaboratif renforcé. Par ailleurs, le web s'étant radicalement transformé en quelques années par l'émergence rapide de nouvelles avancées technologiques et de nouvelles interfaces interactives (appelées « web 2.0 »), le site initial ne correspondait plus aux attentes et exigences des utilisateurs, qu'ils soient simples visiteurs ou professionnels. De plus, ces avancées permettaient d'envisager autrement le fonctionnement de la plateforme, notamment concernant le mode de partage des données entre établissements partenaires du projet, l'association et l'implication de ces partenaires représentant une synergie indispensable à son développement et à son rayonnement.

L'évolution devait nécessairement porter sur la transformation du système de distribution de données précédent, composé d'une vitrine unique, en une plateforme centralisée universelle génératrice d'une constellation de sites satellites orientés par leur contenu et leur cible vers un public précis. Un des objectifs principaux restait celui d'accueillir des documents et objets patrimoniaux de types très variés, conservés dans des établissements très différents entre eux. Fruit de cette réflexion autour de la mise en commun du patrimoine documentaire et de la collaboration entre établissements partenaires, la bibliothèque numérique e-corpus a vu le jour en octobre 2009⁷. Aujourd'hui, e-corpus

regroupe plus de 250 établissements contributeurs venant de 26 pays. Le corpus numérique géré représente à ce jour 1100 000 fichiers numériques (pages de livres, de manuscrits, de journaux ou photographies isolées...) équivalant à plus de 60 000 notices.

Une plateforme collaborative de publication

Le format commun des données descriptives fut établi dans la phase préliminaire (Internum) et il fut décidé de travailler au format EAD XML avec des passerelles simplifiées vers Unimarc et Dublin Core⁸. Le problème n'était donc plus sur ces points mais sur la mise en place pour cette bibliothèque numérique d'une nouvelle interface ayant pour objectif de valoriser la diversité culturelle dans le monde et, à l'époque, bien que cela soit désormais de moins en moins le cas, dans l'espace euro-méditerranéen spécifiquement.

Afin d'assurer son développement, cette plateforme a évolué vers une forme collaborative de publication dans laquelle toute collection peut être mise en valeur par un site internet dédié ayant une adresse internet différente de la base de données racine e-corpus. Il peut s'agir :

- de sites thématiques, regroupant plusieurs établissements et plusieurs collections autour d'un même sujet : citons par exemple la bibliothèque numérique « Camera Obscura » sur l'histoire des techniques photographiques⁹;
- de sites de collectivités convaincues des atouts d'une coopération lo-

cale interétablissements ; ainsi, la ville de Nîmes a décidé de mettre sur une plateforme commune les documents numérisés de la bibliothèque municipale et du musée du vieux Nîmes¹⁰. On y retrouve, dans un corpus virtuellement réuni, des manuscrits, des imprimés de la bibliothèque, des textiles provençaux (les « indiennes » et les châles aux célèbres motifs de cachemire du musée) ainsi que, prochainement, d'anciennes archives industrielles sur la manufacture des textiles régionaux conservées au musée ;

- plus intéressant encore, e-corpus propose aussi des collections virtuelles regroupant des documents d'intérêt commun autour d'une même thématique. Ainsi, les objets du musée du vieux Nîmes sont visibles en perspective avec ceux du musée des traditions provençales de Château-Gombert (Marseille), et même avec d'autres musées d'Inde et du Portugal. Ces deux dernières collections sont elles-mêmes rapprochées, via la collection virtuelle « Eurindia », avec les plans des collections indiennes des Archives nationales d'outre-mer (Aix-en-Provence) et les collections du colonel Gentil, elles-mêmes partagées entre la Bibliothèque nationale de France et le Victoria and Albert Museum. Ainsi, on découvre la force, non pas de l'accompagnement excessif du cyber-usager, mais des mises en relation et des renvois astucieusement organisés pour voyager de collection en collection, de ville en ville, de pays en pays, entre archives, bibliothèques et musées.

Une plateforme de diffusion accessible à tous

L'accès aux collections et aux documents est un élément essentiel de la mise en valeur du patrimoine. Faut-il rappeler que la numérisation en soi est inutile si elle n'est pas assortie d'une stratégie de valorisation et de diffusion ? Cette stratégie permet, dans une certaine mesure, la conservation des documents (sans minorer

dans cette tâche, notamment par la rédaction du nouveau cahier des charges du projet dans le cadre des projets tutorés de leur cursus d'études.

8. EAD XML, Unimarc et Dublin Core sont des schémas de métadonnées génériques qui permettent de décrire des ressources numériques ou physiques et d'établir des relations avec d'autres ressources ; voir notamment : www.archivesdefrance.culture.gouv.fr/gener/classement/normes-outils/ead et <http://dublincore.org>

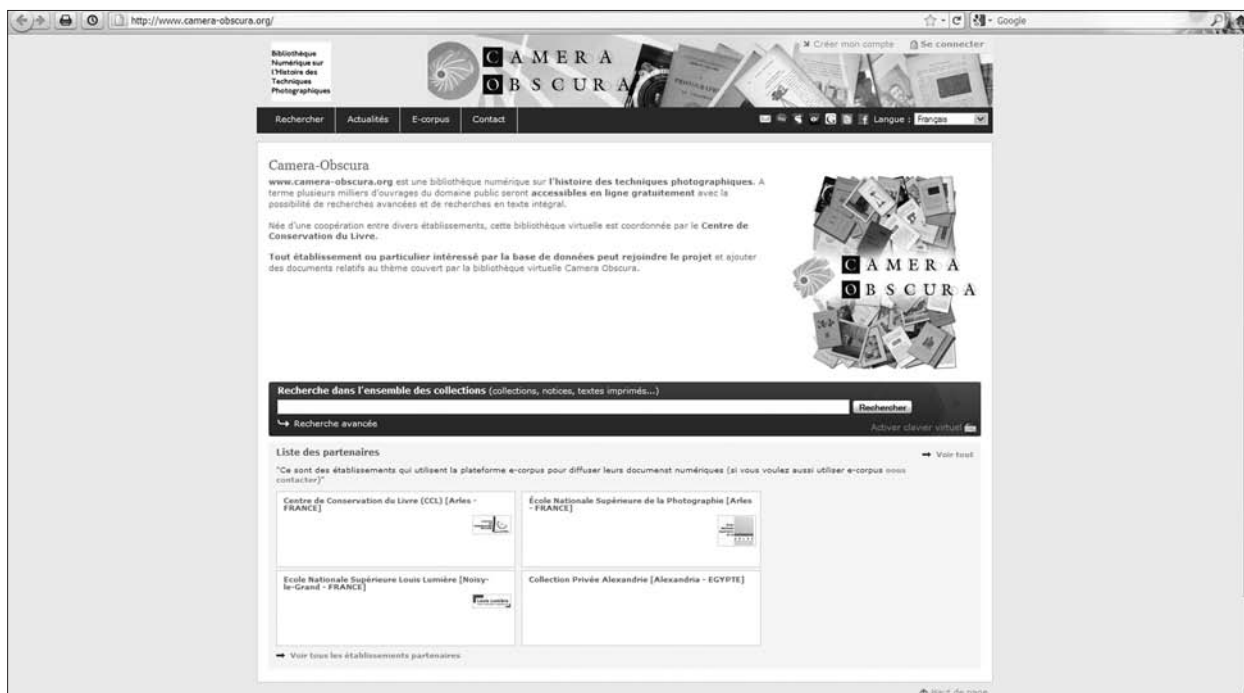
9. www.camera-obscura.org

10. <http://bibliotheque-numerique.nimes.fr>

6. www.unicode.org

7. Saluons à cette occasion l'aide inestimable d'une équipe de jeunes conservateurs stagiaires de l'Enssib : François Mistral, Julie Monier, Sébastien Peyrard, Laure Rioust et Julien Starck n'ont pas ménagé leurs efforts

→ Quand des établissements de conservation du patrimoine mobilier se retrouvent sur www.e-corpus.org :



Page d'accueil de la bibliothèque numérique de Camera Obscura

la très bonne conservation de certains supports de l'écrit, à l'instar des manuscrits byzantins sur parchemin, qui se conservent très bien sans nécessité de reproduction de sauvegarde). Cette numérisation est pourtant nécessaire pour permettre aux citoyens d'avoir un accès simple et gratuit aux documents qu'il n'est pas possible de trop exposer ou de donner en consultation sans contrôle ou limitation.

Pour e-corpus, des outils informatiques performants ont été mis en place afin d'offrir un accès et une consultation plus agréable et ergonomique au public, ainsi qu'une utilisation simplifiée pour les établissements partenaires. Cette nouvelle configuration de la bibliothèque numérique permet :

- la « démocratisation » de l'accès à l'importante banque de données de notices et de documents numérisés que représente e-corpus ;
- la mise en valeur de certaines collections qui, au lieu de se noyer dans la masse d'informations que représente la plateforme aujourd'hui, sont directement accessibles sur des sites dédiés ;
- la valorisation par des partenaires des sous-ensembles de la base

de données par la réalisation de bibliothèques thématiques ou d'expositions virtuelles, au moyen d'outils d'administration de site et de gestion directe de contenus, mais aussi, quand cela est nécessaire, l'aide à la visibilité d'un établissement partenaire, comme on le verra plus loin ;

- l'émergence d'une communauté en ligne qui participe directement ou indirectement au rayonnement du patrimoine documentaire ;
- une minimisation drastique des frais, des tâches et du travail humain par la maintenance et l'administration centralisée et externalisée d'une base de données unique.

L'importance de l'interopérabilité

On l'a dit, les documents sont décrits selon le format XML-EAD, cette standardisation permettant une interopérabilité des informations entre tous les établissements partenaires. L'interface respecte la hiérarchie des classements archivistiques et bibliothéconomiques, ce qui assure l'homogénéité entre l'ensemble des catalogues déjà existants.

La plateforme e-corpus se veut également multilingue. L'interface est ainsi traduite dans une multitude de langues, n'excluant de ce fait aucun musée, bibliothèque et/ou archive susceptibles d'être intéressés par le projet : français, anglais, chinois, arabe, espagnol, catalan, italien (ainsi que russe, géorgien et allemand en cours de réalisation). Une interface de gestion et de traduction des termes utilisés dans le site est intégrée dans le module informatique d'administration de la bibliothèque numérique, ces termes, utilisés notamment pour la recherche de notices, devant avoir obligatoirement une correspondance dans toutes les langues.

Une recherche et une exploitation facile des résultats sont des critères essentiels pour le succès de toute bibliothèque numérique. E-corpus a donc veillé à améliorer l'accès au contenu de la base de données par diverses fonctionnalités :

- reconnaissance optique des caractères automatisée et appliquée sur chaque nouvelle page ou ouvrage intégré dans la base de données, ainsi que sur tout le corpus de documents déjà présents. Ce processus permet aux utilisateurs d'effectuer des recherches de

termes spécifiques directement dans le contenu de l'ouvrage, avec un affichage des résultats hiérarchisé selon un algorithme affinant la pertinence des résultats;

- possibilité pour chaque internaute de créer son propre catalogue (il suffit de créer un compte personnel gratuit) et de l'exporter en format PDF ou EAD;
- ergonomie de la partie administration de la base de données corrigée permettant à chaque partenaire le souhaitant de saisir à distance les informations sur les collections documentaires de son établissement et faisant place à de nouvelles fonctionnalités (création de gabarit CMS, suivi des statistiques pour chaque établissement, etc.);
- visionnage des documents amélioré, avec la possibilité de feuilleter les ouvrages et d'imprimer directement en grand format la ou les page(s) désirée(s);
- amélioration générale de l'interface graphique du site se voulant plus attrayante.

La plateforme permet également l'accès à du contenu hors base de données e-corpus, un « moissonneur » OAI-PMH¹¹ ayant été configuré sur le serveur afin de récolter sur la base de données les métadonnées depuis les entrepôts désignés (à savoir Gallica¹², Europeana). Une option sur le moteur de recherche permet aux utilisateurs de choisir de chercher non seulement sur e-corpus mais également sur l'intégralité des données récoltées sur les entrepôts moissonnés. Le protocole OAI-PMH étant aujourd'hui de plus en plus utilisé à travers le monde par de nombreux acteurs des bibliothèques numériques, e-corpus a fait le choix de favoriser sa coopération avec des sites ayant une interface similaire. Ce protocole simple et interopérable pour le partage des données favorise en effet et par nature la collaboration entre les établissements concernés.

Des partenaires qui participent au développement de la plateforme et aux évolutions

La force de e-corpus, au-delà de ses performances technologiques, est avant tout d'être à l'écoute des besoins des établissements, non pas de manière globale et dirigiste mais de manière individuelle. L'objectif de cette plateforme est de s'adapter à chacun plutôt que de demander aux bibliothèques de se plier à ses exigences. Pour autant, c'est encore souvent le cas, et les grands établissements et grands outils électroniques sont autant de craintes pour une profession qui est en train de prendre conscience qu'elle est l'un des acteurs principaux de la société de l'information, tout en souhaitant que cela ne soit pas à n'importe quel prix et au détriment des acquis de méthodes professionnelles éprouvées laissant l'humain au centre du dispositif.

De nombreux chantiers restent à mener à bien, comme la consultation de la presse, qui devrait se démarquer d'autres corpus dans sa présentation (voir notamment la collection des périodiques de la bibliothèque Méjanes, *Le Mémorial d'Aix*), à l'instar de la proposition de la bibliothèque municipale de Lyon dans le domaine¹³.

Savoir qu'un nouveau projet de description, demandant l'intégration de nouveaux champs de description, peut être mis en place sans efforts et sans qu'il vous soit répondu qu'il fallait y penser dans la phase initiale du projet – ce qui arrive trop souvent pour les bases de données développées par les chercheurs pour un projet spécifique sans prise en compte des questions d'interopérabilité et de pérennisation des données – permet de réfléchir à de nouveaux projets de manière apaisée.

Un espace ouvert et collaboratif

Les bibliothèques, les musées et les archives peuvent exploiter gratuitement e-corpus pour la diffusion de leurs fonds numérisés et profiter ainsi des outils informatiques disponibles sur la plateforme. L'hébergement et la diffusion sont gratuits, de même que la création de sites dédiés que les établissements concernés peuvent gérer de manière autonome : la véritable nouveauté réside dans cette autonomie offerte aux établissements intéressés.

Plusieurs établissements peuvent également se regrouper pour créer et animer de manière collaborative une bibliothèque numérique, ou même simplement un catalogue collectif consacré à un sujet précis. En effet, la bibliothèque numérique bénéficie d'un site principal global et permet, à partir d'une base de données centralisée, de créer toute une constellation de sites dits satellites. Chacun d'entre eux a sa propre adresse internet et fonctionne en tant qu'entité à part entière, distincte d'e-corpus, référencée en tant que telle dans les moteurs de recherche et annuaires internet, et adoptant sa propre identité visuelle. Grâce à la centralisation des informations, toute opération (saisie de notices, import de documents numériques, etc.) est immédiatement et globalement effective dans tous les sites satellites, ce qui supprime toute redondance des tâches et réduit considérablement les coûts de fonctionnement et d'exploitation.

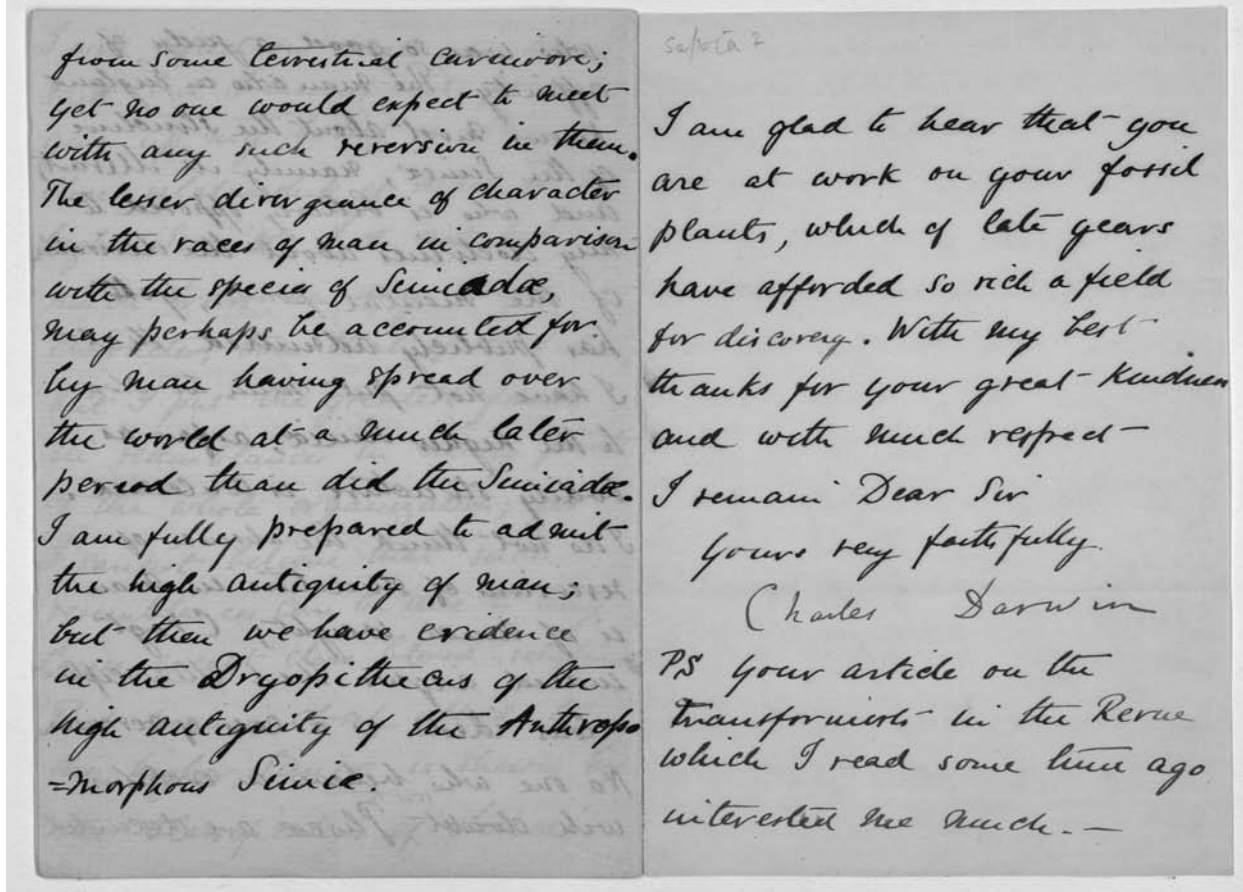
Ce type d'outil et plus largement la participation des bibliothèques, des archives et des musées au projet e-corpus a pour avantage :

- d'améliorer l'interopérabilité, l'intégration et la transparence de l'accès aux ressources des établissements;
- de réduire les coûts pour les musées, les archives et les bibliothèques qui désirent partager leurs contenus numériques avec le public;
- d'accroître la capacité des musées, des archives et des bibliothèques de partager à vaste échelle du contenu numérique en utilisant des normes techniques et descriptives de façon pratique et économique et également de partager des outils communs;

11. Open Archives Initiative Protocol for Metadata Harvesting; voir : www.openarchives.org/OAI/openarchivesprotocol.html

12. <http://gallica.bnf.fr>

13. La bibliothèque municipale de Lyon a entrepris la numérisation de sa collection de journaux lyonnais parus entre 1830 et 1914 et a développé une plateforme dédiée à ce corpus : <http://collections.bm-lyon.fr/PER003>



Extrait d'une lettre de Charles Darwin (Angleterre) écrite à Gaston de Saporta (France) – 1872. Archives Gaston de Saporta.
Document disponible sur e-corpus

- d'offrir aux internautes un accès facilité et unifié à de très nombreuses collections numériques, sans avoir à consulter des dizaines de bibliothèques numériques dans l'espoir fou de trouver peut-être tel ou tel document qu'ils souhaitent consulter.

De plus, la création de collections virtuelles thématiques a pour avantage d'accroître la visibilité et l'accessibilité de tel ou tel sujet auprès du public et incite les bibliothèques, les archives et les musées à partager leurs ressources documentaires et iconographiques.

Plusieurs établissements se sont regroupés, par exemple, autour d'un thème commun, la Provence, dans la Bibliothèque provençale numérique, lancée conjointement par le Centre de conservation du livre et la ville d'Aix-en-Provence. Elle propose, présente et met en ligne un ensemble de documents patrimoniaux sur la Provence et son aire culturelle, conservés dans des bibliothèques, des archives, des institutions muséales ou des collections privées, que ce soit dans la Région Provence-Alpes-Côte d'Azur, en France et à l'étranger : livres, images, manuscrits, journaux, mais aussi enregistrements sonores et vidéos – tous les documents sont libres de droits. Exemple de synergie entre musées, bibliothèques et archives, ce site satellite permet le

regroupement d'informations au sein d'un portail commun et une gestion partagée de son contenu directement par les établissements concernés. Autrement dit, tout en favorisant la mise en commun de ressources, e-corpus accroît le travail collaboratif entre les professionnels. Ce type de collaboration permet également de composer de nouveaux auditoires, par croisement, pour tous les partenaires.

Conclusion

À l'interface des archives, des bibliothèques et des musées se trouve désormais internet. Il faut que nous le prenions comme une chance et un nouveau défi pour nos établissements. L'exemple d'e-corpus montre le chemin parcouru en une quinzaine d'années, le potentiel de l'outil informatique et du réseau internet ainsi que les enjeux à relever dans la décennie à venir. E-corpus n'est pas le seul projet de ce type, le mouvement est lancé depuis longtemps. Les exemples existants fonctionnent, qu'ils soient basés sur la collaboration de bibliothèques numériques déjà constituées, dont la consultation est centralisée, ou sur le modèle de la base collective. Ils fonctionnent au niveau international,

national ou régional : Europeana, Gallica, la Banque numérique du savoir d'Aquitaine (BnsA)¹⁴ en sont de très bons exemples. Les projets thématiques, souvent très pointus, sont parfois plus difficiles à faire connaître et s'adressent à un public souvent très réduit (Roman de la Rose). Pourtant, on note une résistance encore grande de nombreux établissements, trop souvent attachés à créer leur propre bibliothèque numérique en faisant ainsi fi de la manière dont les internautes utilisent internet¹⁵. Que faut-il encore faire pour convaincre bibliothécaires, archivistes, responsables de musées et chercheurs de s'unir et travailler ensemble ? ●

Mai 2011

14. La BnsA est un programme éditorial numérique organisant la conception, l'organisation, la production, l'édition et la mise à disposition en ligne de données numériques relatives au patrimoine aquitain. Voir dans ce numéro du BBF l'article de Jean-François Sibers, « Banque numérique du savoir en Aquitaine : dix ans pour la confluence des ressources patrimoniales en région », p. 50-56.

15. Isabelle Westeel, « Le patrimoine passe au numérique », BBF, 2009, n° 1, p. 28-35. En ligne : <http://bbf.enssib.fr/consulter/bbf-2009-01-0028-003>