



La bibliothèque numérique américaine est lancée !

ROBERT DARNTON

Cet article est initialement paru dans la New York Review of Books le 25 avril 2013. Il est consultable en ligne dans sa version originale :

www.nybooks.com/articles/archives/2013/apr/25/national-digital-public-library-launched/?pagination=false

Nous remercions la revue pour son autorisation gracieuse de proposer aux lecteurs du BBF une traduction en français. L'article a été traduit par Emmanuel Jaslier.

Robert Darnton, universitaire, est un spécialiste de l'Europe moderne et de l'histoire du livre. Il a enseigné l'histoire de l'Europe à l'université de Princeton (1968-2007), et est directeur général des bibliothèques de l'université de Harvard, et professeur titulaire de la chaire d'histoire moderne Carl H. Pforzheimer. Il a publié de nombreux ouvrages sur le siècle des Lumières en France et sur l'histoire du livre : *L'aventure de l'Encyclopédie : 1775-1800 : un best-seller au siècle des Lumières*, préface d'Emmanuel Le Roy Ladurie (Paris, Perrin, 1982), ou encore *Apologie du livre : demain, aujourd'hui, hier* (Paris, Gallimard, 2010).

La Bibliothèque publique numérique d'Amérique (DPLA, Digital Public Library of America), lancée le 18 avril 2013, est un projet visant à rendre accessibles en ligne à l'ensemble des Américains – et, à terme, au monde entier – les fonds des bibliothèques de recherche, des centres d'archives et des musées des États-Unis, et ce, gratuitement. Comment est-ce possible ? Pour répondre à cette question, je vais évoquer les premiers pas de la DPLA, ainsi que son avenir immédiat. Mais avant tout, je pense nécessaire de prendre du recul pour apprécier la façon dont un projet aussi ambitieux participe au développement de ce qu'on appelle communément la société de l'information.

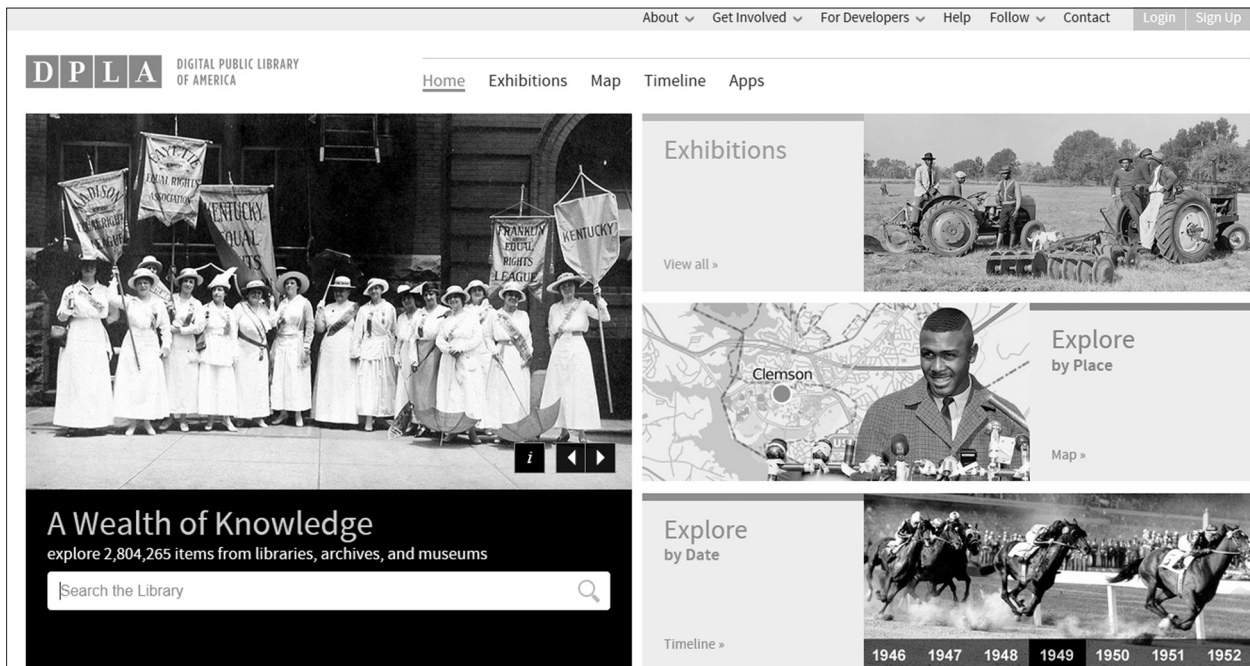
Pour faire simple, la DPLA résulte de la confluence de deux courants qui ont façonné la civilisation américaine : l'utopisme et le pragmatisme. La tendance utopique marque la République à sa naissance, car les États-Unis ont été créés par une révolution et les révolutions sont sources d'énergie utopique – à savoir, la conviction que les choses ne sont pas telles qu'elles devraient être. Quand tout s'effondre, avec violence et par l'action collective, il devient possible de tout reconstruire, en se fondant sur des principes meilleurs.

Les révolutionnaires américains s'inspiraient des Lumières, ainsi que d'autres sources, parmi lesquelles des variantes hétérodoxes de l'expérience religieuse et la conviction bien ancrée que leur naissance en tant qu'Anglais libres leur conférait certains droits inaliénables. Mélangez bien ces ingrédients, et vous obtenez la Déclaration d'indépendance et la Déclaration des droits – des affirmations de principes radicales qui ne passeraient jamais au Congrès de nos jours.

Cependant, les révolutionnaires américains étaient des hommes pragmatiques qui avaient un travail à accomplir. Lorsque les Articles de la Confédération s'avèrent inadaptés, ces hommes s'attachèrent à édifier une union plus parfaite et se remirent au travail, aboutissant à une Constitution visant à rendre plus efficace l'État en lui accordant plus de pouvoir, tout en lui imposant des organes de contrôle. Création de freins et de contrepoids, rédaction du *Fédéraliste*, empoignades viriles au sujet de la course vers la richesse et le pouvoir, sans même parler de débats sur l'esclavage et les salaires de misère : les pères fondateurs étaient des durs et âpres en affaires !

Comment ces deux tendances convergent-elles dans la Bibliothèque publique numérique d'Amérique ? Nonobstant la technologie futuriste sur laquelle elle s'appuie, la DPLA remonte au dix-huitième siècle. Quoi de plus utopiste qu'un projet ambitionnant de rendre l'héritage culturel de l'humanité accessible à tout être humain ? Quoi de plus pragmatique qu'un système agrégeant des millions de mégabits afin de les proposer aux lecteurs sous la forme de textes facilement accessibles ?

Par-dessus tout, la DPLA est l'expression de la foi des Lumières dans le pouvoir de la communication. Jefferson et Franklin – le chantre de la Bibliothèque du Congrès et l'imprimeur devenu philosophe-homme d'État – partageaient la conviction profonde que la santé de la République dépendait de la libre circulation des idées. Ils savaient que la diffusion de la pensée était tributaire de l'imprimerie. Toutefois, la technologie de l'imprimerie n'avait que peu évolué depuis l'époque de Gutenberg et n'était pas



Détail de la page d'accueil de la DPLA qui peut être consultée à l'adresse <http://dp.la>

un vecteur suffisamment puissant pour répandre la bonne parole dans toutes les couches d'une société faiblement alphabétisée et présentant un fort degré de pauvreté.

Grâce à internet et à un système d'éducation omniprésent à défaut d'être parfait, nous pouvons maintenant réaliser le rêve de Jefferson et de Franklin. Nous disposons des ressources technologiques et économiques suffisantes pour rendre accessible l'intégralité des collections de toutes nos bibliothèques à l'ensemble de nos concitoyens – ainsi qu'à n'importe qui, n'importe où, s'il dispose d'un accès au World Wide Web. Telle est la mission de la DPLA.

Présenté aussi audacieusement, ce projet semble démesuré. On peut facilement se laisser entraîner par la rhétorique utopiste sur la bibliothèque des bibliothèques, la mère de toute bibliothèque, la bibliothèque d'Alexandrie des temps nouveaux. Pour créer la DPLA, nous avons dû recourir à la détermination, à l'expérience de terrain, au pragmatisme professionnel ancré dans la tradition américaine. Je vais maintenant décrire ce qu'est la DPLA, ce qu'elle va offrir au public américain au moment de sa mise en service, et ce qu'elle est appelée à devenir dans un avenir proche.

Comment la conceptualiser ? Ne pensez pas à une grande structure pyramidale fondée sur une base de données gigantesque. La DPLA sera

un système distribué de contenus électroniques rendant accessibles, gratuitement et sans effort, à tout lecteur connecté sur le web, les fonds des bibliothèques publiques et de recherche, des archives, des musées et des sociétés d'histoire. Pour que ça marche, il est nécessaire de penser grand et de commencer petit. Au début, l'offre de la DPLA sera limitée à une riche sélection de collections – livres, manuscrits et œuvres d'art – qui ont déjà été numérisées par des institutions culturelles dans tout le pays. Le corpus grandira à partir de ce noyau, accumulant progressivement des ressources de tout type jusqu'à constituer une bibliothèque numérique nationale.

La raison de cette stratégie de développement est à chercher dans l'histoire de son origine – car la DPLA a une histoire, même si elle n'a pas trois ans d'âge. Elle a vu le jour lors d'une conférence à Harvard tenue le 1^{er} octobre 2010, un événement discret réunissant une quarantaine de personnes, pour la plupart des responsables de fondations ou de bibliothèques. Dans la lettre d'invitation, j'avais inclus une note d'une page centrée sur la proposition suivante : « rendre la masse de la littérature mondiale accessible à tous les citoyens gratuitement » en créant « une grande coalition des fondations et des bibliothèques de recherche ». Rétrospectivement, cela ressemblait fort à une utopie, mais tous au cours de ce séminaire se sont

accordés pour dire qu'un tel travail valait le coup d'être fait et qu'on avait les moyens de le faire.

Nous sommes également tombés d'accord sur une courte description qui est devenue aujourd'hui un énoncé de mission. La DPLA, a-t-il été décidé, serait « un réseau ouvert et distribué de ressources complètes en ligne qui s'appuierait sur le patrimoine national des bibliothèques, des universités, des centres d'archives et des musées afin d'éduquer, d'informer et de renforcer l'autonomie de chacun pour les générations présentes et à venir ».

Ça sonne bien, me direz-vous, mais Google ne fournit-il pas déjà un tel service ? C'est exact, Google avait démarré courageusement une entreprise de numérisation de tous les livres dans le monde et a réussi à créer une base de données gigantesque qui, au dernier comptage, regroupait trente millions de volumes. Mais ce projet s'est heurté aux lois sur le droit d'auteur et à des procès intentés par les détenteurs de ces droits. Google a essayé de rallier les parties adverses en les invitant à s'associer en tant que partenaires dans un projet encore plus grand. Ils ont trouvé un accord qui a transformé l'entreprise initiale de Google, un moteur de recherche qui n'affichait que de courts extraits des livres, en une bibliothèque commerciale. En souscrivant à un abonnement, les bibliothèques de recherche accédaient à la base de données de

Google – c'est-à-dire, aux versions numérisées des livres qu'elles avaient elles-mêmes fournis à Google gratuitement et qu'elles pouvaient maintenant rendre accessibles à leurs usagers au prix fixé par Google et ses nouveaux partenaires. Pour certains d'entre nous, Google Book Search constituait alors un nouveau monopole de l'accès au savoir. Pour le tribunal fédéral du district sud de New York, cet accord était tellement truffé de dispositions inacceptables qu'il ne pouvait exister au regard de la loi.

Suite à la décision de justice du 23 mars 2011¹ rejetant cet accord, la bibliothèque numérique de Google était bel et bien morte, même si Google peut continuer à utiliser sa base de données pour d'autres usages, comme la fourniture commerciale de reproductions numériques en accord avec les éditeurs concernés. La DPLA n'a pas été conçue pour remplacer Google Book Search; en fait, sa conception remonte bien avant cette décision de justice, mais la DPLA s'est inspirée de cette audacieuse tentative de Google de numériser des fonds entiers issus des bibliothèques et n'a pas renoncé à l'espoir de convaincre Google de s'allier à elle pour travailler pour le bien public.

On pourrait toutefois élever une autre objection : qui a autorisé ce groupe auto-constitué à s'attaquer à un tel projet ?

Réponse : personne. Nous sommes persuadés qu'une initiative privée était nécessaire et que le projet n'aurait jamais démarré si nous avions attendu que le gouvernement agisse. Par conséquent, nous avons nommé un comité directeur, un secrétariat installé

au centre Berkman à Harvard et six groupes répartis dans tout le pays, qui ont commencé à étudier et à débattre des questions clés : pilotage, financement, infrastructure technologique, droit d'auteur, périmètre et contenu des collections, ainsi que le public auquel elles seraient destinées.

Les groupes se sont développés et ont pris leur propre essor, s'appuyant sur des contributions bénévoles, du *crowdsourcing* (pratique consistant à faire appel aux contributions d'un groupe indéfini, généralement une communauté en ligne, comme le fait par exemple Wikipedia) et des discussions par le biais de sites web, de listes de diffusion, de réunions publiques et de séminaires extrêmement spécialisés. Des centaines de personnes se sont engagées activement et des milliers d'autres ont participé au débat sans fin et bruyant mené sur internet. Des rencontres plénières à Washington DC, San Francisco et Chicago ont attiré des foules conséquentes et un public virtuel encore plus nombreux, grâce aux textos, aux tweets, au streaming et autres formes de échanges. S'est ainsi forgé un sens de la communauté, façon vingt et unième siècle : ouvert, brouillon, virtuel, mais néanmoins réel, car soutenu par une ossature édifiée dans le web.

Le débat, aussi bien réel que virtuel, a pris place alors que les groupes se sont mis au travail. Quarante volontaires ont soumis des versions bêta des prototypes de logiciel que la DPLA pourrait utiliser, soumis à un « bêta-test », une forme d'évaluation conduite par les usagers. Après de nombreux tests et remaniements, une plate-forme a été développée afin de fournir les liens vers le contenu des collections de bibliothèque dans tout le pays et d'agrèger les métadonnées – c'est-à-dire, les informations de type catalographique qui identifient les fichiers numériques et décrivent leur contenu. Les métadonnées seront agrégées dans un réservoir situé dans ce que les concepteurs appellent la dorsale (*backend*) de la plate-forme, et une interface de programmation (API, Application Programming Interface) en « frontal » permettra à tout type de logiciels de transmettre de diverses manières les contenus aux usagers.

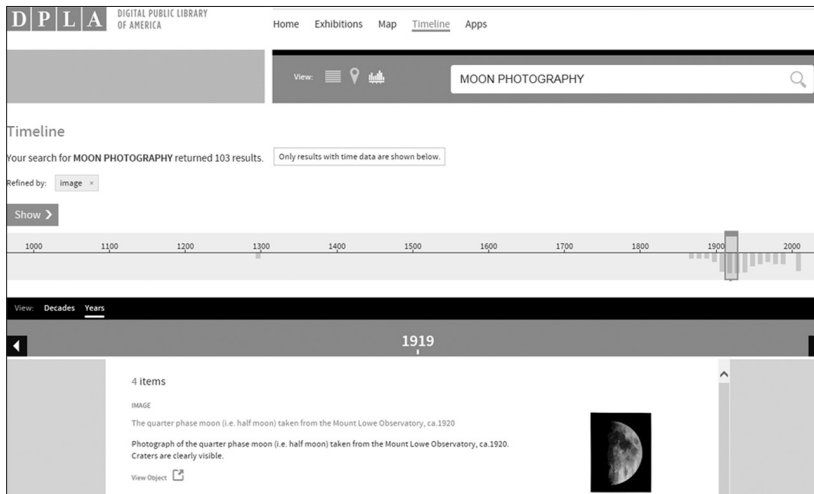
L'interface conviviale permettra ainsi à n'importe quel lecteur – par exemple, un lycéen du Bronx – de consulter des documents jusqu'ici stockés dans des rayonnages inaccessibles ou conservés dans des réserves aux accès limités – par exemple, les pamphlets traitant de l'invalidation et de la sécession dans le Sud d'avant la guerre civile conservés à la bibliothèque de Huntington à Los Angeles. Les usagers se connecteront simplement à la DPLA par son adresse URL, <http://dp.la>. Ils pourront alors rechercher des documents par leur titre ou le nom d'un des auteurs et auront accès via le site de la DPLA au livre ou à tout autre type de document numérique, directement sur le site de son institution d'hébergement. L'illustration de la page précédente montre à quoi ressemble l'écran de l'utilisateur, bien qu'il s'agisse encore d'une maquette test.

Parallèlement, nombre des plus grands musées et bibliothèques du pays – dont Harvard, la New York Public Library et le Smithsonian – sont prêts à rendre accessible au public une sélection de leurs collections via la DPLA. Ces documents seront consultables par tous dès le lancement du 18 avril, mais ils ne constituent que le début de cette offre agrégative qui continuera à croître organiquement autant que les budgets et les lois sur le droit d'auteur le permettront.

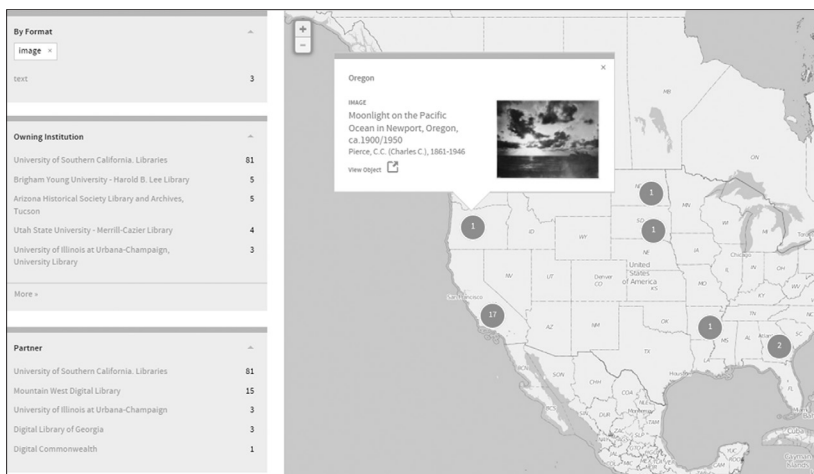
Bien sûr, cette croissance doit être durable. Mais les plus importantes fondations du pays ont exprimé de l'intérêt pour le projet. Plusieurs d'entre elles – les fondations Sloan, Arcadia, Knight et Soros, ainsi que le National Endowment for the Humanities et l'Institute of Museum and Library Services – ont financé les trois premières années d'existence de la DPLA. Si une douzaine de fondations combinaient leurs moyens, assignant une somme fixe à un budget annuel, nous pourrions créer en une décennie l'équivalent numérique de la Bibliothèque du Congrès. Les mécènes espèrent bien naturellement que la Bibliothèque du Congrès participera également à la DPLA.

L'obstacle principal à la croissance de la DPLA n'est pas d'ordre financier, mais juridique. Les droits d'auteur peuvent exclure du corpus tout ce qui

1. Voir : Robert Darnton, « Google's Loss : The Public's Gain », *The New York Review of Books*, 28 avril 2011. En ligne : www.nybooks.com/articles/archives/2011/apr/28/google-loss-publics-gain/?pagination=false
 NDLR : lire la chronologie des faits dans l'article de Miguel Helft, « Judge Rejects Google's Deal to Digitize Books », *New York Times*, 22 mars 2011. En ligne : www.nytimes.com/2011/03/23/technology/23google.html?_r=0
 – Dépêche de l'AFP, « La justice américaine rejette l'accord entre Google et les éditeurs », 22 mars 2011, à lire sur le site du *Monde* : www.lemonde.fr/technologies/article/2011/03/22/la-justice-americaine-rejette-l-accord-entre-google-et-les-editeurs_1497074_651865.html



Visualisation au travers d'une frise chronologique des réponses à une recherche sur « moon photography ».



Géolocalisation des bibliothèques détentrices des fonds.

a été publié après 1964, la plupart des documents publiés après 1923, et même certains dont la date de publication peut remonter jusqu'à 1873. Cependant, des jugements rendus ces derniers mois permettent désormais d'étendre les dispositions sur l'usage loyal (*fair use*) établies par la loi sur le droit d'auteur de 1976 à des ouvrages récents dans le cadre de certains types d'utilisation, comme par exemple les services aux déficients visuels ou certains services pédagogiques. Et si, comme prévu, la DPLA exclut d'emblée les ouvrages encore en vente sur le marché (la plupart n'ont une durée de vie commerciale que de quelques

années), il est possible que des auteurs et des éditeurs acceptent de lui céder leurs droits.

Quoi qu'il en soit, il valait mieux ne pas attendre que les tribunaux démêlent cet écheveau juridique pour créer une structure administrative efficace. Le secrétariat officiel à Harvard a été remplacé par une association à but non lucratif organisée suivant les recommandations du Code américain des impôts. Un conseil d'administration a succédé au comité de pilotage. Et les six groupes vont évoluer en un système de comités aux attributions précisément définies, comme la sensibilisation auprès des bibliothèques

publiques et des *community colleges*². La nomination en tant que directeur exécutif de Daniel Cohen de l'université George Mason, grand historien et expert de l'internet, a été annoncée le 5 mars; les premiers membres de l'équipe ont déjà été engagés; et les bureaux de l'administration vont être installés à Boston.

Ces premières étapes ne conduiront pas à la création d'une bureaucratie fortement centralisée. Au contraire, l'aspect « distribué » de la DPLA garantit que son action sera répartie à travers tout le pays. Son réservoir sans cesse grandissant de métadonnées (Harvard a déjà rendu accessible douze millions de notices de métadonnées) sera stocké dans des nuages informatiques (*computer clouds*), et ses activités seront canalisées dans deux types de « hubs ».

Les « hubs de contenu » de la DPLA sont de larges réservoirs de données numériques, généralement stockés dans des lieux comme l'Internet Archive de San Francisco. Ces données seront accessibles aux usagers directement via la DPLA sans passer par un agrégateur intermédiaire. Les « hubs de service » – des centres de collecte de contenu – agrégeront des données et fourniront des services variés à l'échelle régionale ou nationale. La DPLA ne peut travailler directement avec toutes les bibliothèques, tous les centres d'archives et tous les musées des États-Unis, sous peine de voir son administration centrale accaparée par le développement de centaines de milliers d'interfaces et de liens. Mais le développement en la matière au niveau local est désormais coordonné au niveau fédéral et la DPLA va travailler avec les États en vue de créer un système intégré national.

Quarante États disposent d'une bibliothèque numérique. Les hubs de service de la DPLA – sept sont déjà en cours de développement en différents endroits du pays – exploiteront les données numériques qu'elles ont

2. NDLR : ces collèges communautaires (publics) dispensent un enseignement pour les deux premières années des quatre années d'université permettant d'obtenir un certain type de diplômes ou certificats de niveau intermédiaire.

déjà collectées pour le réseau national. Entre autres fonctions, ces hubs de service aideront les bibliothèques locales et les sociétés d'histoire à scanner, mettre en valeur et préserver les documents locaux – sources de la guerre de Sécession, annuaires de lycées, correspondances personnelles, tout ce que leurs collections recèlent ou que leurs usagers voudront exhumer de leur malle ou de leur grenier. En se développant, l'autonomisation numérique à l'échelon le plus local participera à la construction d'une collection intégrée au niveau national, et cette collection nationale sera liée à celles d'autres pays.

La DPLA a conçu son infrastructure de sorte qu'elle soit interoperable avec Europeana, un super-agrégateur soutenu par l'Union européenne, qui coordonne les opérations de liens entre les collections de vingt-sept pays européens. En une génération, il devrait s'être constitué un réseau à l'échelle mondiale qui rendra accessible presque toutes les collections des bibliothèques et des musées à presque tout le monde sur la planète. Afin de donner un aperçu de ce que pourrait être ce futur, Europeana et la DPLA ont réalisé une exposition numérique commune sur le thème de l'immigration européenne vers les États-Unis, qui sera consultable en ligne dès le lancement du 18 avril.

Il va de soi que cette expansion, à l'échelon local ou mondial, dépend de la capacité des bibliothèques et des autres institutions à développer leurs propres bases de données numériques – un processus à long terme, irrégulier, qui nécessite un investissement en argent et en énergie. Ce faisant, de grandes réserves de richesses numériques vont se constituer partout sur la planète. Il en existe déjà beaucoup, les plus grandes bibliothèques ayant déjà numérisé des pans entiers de leurs collections, qui constitueront autant de «hubs de contenu» à part entière.

Par exemple, Harvard, en tant que «hub», projette de rendre accessibles à la DPLA dès son lancement 243 manuscrits médiévaux, 5741 pamphlets latino-américains rares, 3628 daguerrotypes ainsi que les premières photographies de la lune ou encore celles d'esclaves d'origine africaine,

502 «chapbooks³» et romans à sensations, un genre littéraire populaire aux dix-huitième et dix-neuvième siècles et 420 récits d'audiences de procès portant sur des affaires de mariage ou de mœurs. Harvard compte fournir bien plus encore dans les mois qui suivront, en particulier dans les thématiques de la musique, de la cartographie, de la zoologie et de l'histoire coloniale. Bien d'autres bibliothèques, centres d'archives et musées contribueront également à cet accroissement grâce à des documents issus de leurs collections. En tout, deux à trois millions de documents de tout format seront accessibles dès le 18 avril.

Quel sera l'usage d'un tel fonds ? Je voudrais conclure par un dernier exemple. Environ quatorze millions d'étudiants luttent pour obtenir un diplôme dans un *community college* – autant que l'effectif intégré dans les collèges classiques et les universités du pays. Mais nombre d'entre eux – et bien plus encore de lycéens – n'ont pas accès à une bibliothèque digne de ce nom. La DPLA peut leur fournir une impressionnante collection numérique et elle peut adapter son offre à leurs besoins. Bien des manuels et des ouvrages de référence dans des domaines comme les mathématiques ou l'agronomie sont toujours d'actualité, même s'ils ne sont plus soumis au droit d'auteur. Remaniés par des experts, ils pourraient être adaptés pour des cours d'initiation et combinés en une bibliothèque de référence pour débutants.

Tôt ou tard, quasiment n'importe quel étudiant trouve sur sa route un poème d'Emily Dickinson, qui est probablement le poète favori des Américains. Mais les poèmes d'Emily Dickinson posent des problèmes particuliers. Seuls quelques-uns ont été publiés, horriblement massacrés, de son vivant. Presque tous les manuscrits sont conservés à la bibliothèque de Houghton à Harvard et ils constituent autant de casse-têtes de par leur ponctuation bizarre, leurs majuscules, leurs espaces et d'autres aspects de leur présentation qui ont des implications profondes sur leur interpré-

tation. Harvard a numérisé les originaux, les a associés à leurs éditions imprimées les plus reconnues (celle publiée par Thomas H. Johnson en 1955 et celle éditée par Ralph W. Franklin en 1981) et y a adjoint une documentation complémentaire pour en faire les «Emily Dickinson Archives», qui seront accessibles par le site de Harvard et la DPLA.

Les archives en ligne enrichiront le travail des étudiants à tous les niveaux de leur cursus. Les enseignants pourront en tirer des sélections de documents adaptés aux besoins de leur classe. En comparant étroitement différentes versions d'un poème, les élèves commenceront à appréhender la façon dont fonctionne la poésie. Ils amélioreront leur maîtrise de la langue en général et les enseignements qu'ils en tireront les aideront à prendre conscience de leur patrimoine culturel. Il s'agit peut-être d'un petit pas, mais il représente une avancée concrète pour le monde du savoir, que Jefferson, dans la droite ligne utopiste, décrivait comme «le bien commun de l'humanité». ●

3. NDLR : petits ouvrages équivalents de la Bibliothèque bleue.