

La Bibliothèque numérique du patrimoine ibéro-américain

JOSÉ LUIS BUEREN GÓMEZ-ACEBO

*Bibliothèque nationale d'Espagne
jose Luis.bueren@bne.es*

José Luis Bueren Gómez-Acebo est responsable du service Bibliothèque numérique au sein de la Direction de la bibliothèque numérique et des systèmes d'information de la Bibliothèque nationale d'Espagne.

Le portail de la BDPI¹ donne accès aux collections numérisées de différentes bibliothèques nationales ibéro-américaines – sept, pour le moment. Le projet, qui a vu le jour dans le cadre de l'Asociación de Bibliotecas Nacionales Iberoamericanas (Association des bibliothèques nationales ibéro-américaines – ABINIA)², vient de fêter sa première année d'activité, et c'est donc le moment idéal pour faire un état des lieux de son fonctionnement actuel et des perspectives futures qui sont les siennes.

Basée sur des standards et des technologies Open Source (SOLR, Java, OAI-PMH, Marc, Dublin Core), la BDPI centralise à l'heure actuelle les ressources numériques des bibliothèques nationales brésilienne, colombienne, chilienne, salvadorienne, espagnole, panaméenne et portugaise, soit plus de 181 000 titres représentant près de 1 255 000 documents numériques.

Pourtant, au-delà de ce qu'elle représente aujourd'hui (un an après sa création, c'est sans nul doute le projet le plus important de la coopération bibliothécaire ibéro-américaine), ce qu'elle pourrait incarner dans un avenir plus ou moins proche est certainement plus significatif encore.

En effet, la BDPI a rapidement occupé une place privilégiée pour en venir à jouer, dans son propre espace géographique, un rôle similaire à celui d'Europeana³ ou de la Digital Public

Library of America (DPLA)⁴. Même si la comparaison peut sembler excessive, s'agissant d'un projet qui est encore loin de pouvoir se mesurer à de tels sites compte tenu des institutions participantes, du nombre d'objets numériques qu'il réunit et du budget dont il dispose, le fait est que la BDPI, tant d'un point de vue conceptuel que technique, diffère assez peu de ses deux grandes sœurs.

On ne peut d'ailleurs nier qu'un vide existait dans ce contexte de projets de coopération internationale : l'Europe, avec The European Library⁵ ou Europeana, et les États-Unis, avec American Memory ou la DPLA, laissaient le champ libre à un portail couvrant l'espace ibéro-américain.

En outre, l'histoire bien spécifique des pays qui composent l'ABINIA contribue à faire de ce portail un outil privilégié pour l'étude historique de chacun de ces pays, en corrélation avec les autres. Quant à l'Espagne, ses relations avec l'Amérique latine constituent une dimension fondamentale de son propre passé et de son identité.

Si la BDPI ne regroupe pour le moment que des bibliothèques nationales, celles qui sont représentées par l'ABINIA, il est tout à fait envisageable, si l'ABINIA en décide ainsi, que d'autres bibliothèques ou même d'autres types d'institutions (archives ou musées) la rejoignent, comme c'est le cas au sein d'Europeana.

La BDPI, depuis le début de sa création, a fait la preuve de sa stabi-

1. Biblioteca Digital del Patrimonio Iberoamericano.

2. www.abinia.org

3. www.europeana.eu

4. <http://dp.la>

5. <http://theeuropeanlibrary.org>

lité et de sa capacité de croissance; elle pourra donc, si des décisions sont prises dans ce sens, contribuer à compléter le panorama international de l'accès à des ressources documentaires numérisées.

Historique du projet

L'idée de la BDPI est née en 2009 lorsque, au sein de l'ABINIA, fut décidée la mise en place d'un portail permettant la consultation depuis une unique page de toutes les collections numériques des bibliothèques nationales ibéro-américaines.

Auparavant, un projet mené entre 2001 et 2003 en avait été une sorte d'ébauche : la bibliothèque numérique El Dorado, créée alors sous l'égide de l'Unesco et de l'ABINIA, avait pour objectif de mettre à la disposition du public des ressources numériques de bibliothèques ibéro-américaines via un protocole Z39.50. Depuis 2003, le contexte de la numérisation a fortement évolué mais indubitablement c'est à ces premiers efforts de coopération et de normalisation que nous devons les réussites actuelles.

Pour revenir aux origines de la BDPI, lorsque sa création fut envisagée en 2009, la Biblioteca Nacional de España (BNE), qui un an plus tôt avait lancé un grand plan de numérisation massive et présenté son nouveau site, Biblioteca Digital Hispánica⁶, s'est proposée pour diriger et coordonner ce nouveau projet.

La BNE a d'abord tenté d'adapter le National Libraries Global Prototype, un *software* conçu pour la bibliothèque nationale de Nouvelle-Zélande (NLNZ), afin de le conformer aux nécessités du projet qui se mettait en place. Il fallait pour cela développer certaines fonctionnalités mais aussi procéder à des transformations ergonomiques.

Parallèlement, la BNE avait envoyé fin 2010 différents questionnaires aux bibliothèques nationales ibéro-américaines potentiellement impliquées pour connaître précisément l'état de leurs bibliothèques numériques et



Une recherche sur le site de la BDH s'exécute simultanément sur Europeana et la BDPI.

envisager leur intégration. Dans le même temps, l'ergonomie du portail était travaillée.

En 2011, après évaluation du prototype, on en vint à la conclusion qu'il offrait des fonctionnalités intéressantes, malgré quelques inconvénients techniques et même stratégiques (comme par exemple le fait que ce prototype n'était installé nulle part ailleurs). S'il était possible de remédier à ces inconvénients, la situation de la BNE elle-même avait bien changé depuis 2009.

Le projet de numérisation qui avait débuté en 2008 à la BNE avait débouché sur la mise en place de la Biblioteca Digital Hispánica (BDH), qui en 2011 avait évolué et témoignait d'une grande solidité et capacité d'adaptation. De plus, elle pouvait supporter toutes les fonctionnalités prévues initialement pour la BDPI, dont l'implantation, la maintenance et le développement futur promettaient d'être plus simples, plus rapides, et moins coûteux, puisque le projet allait de pair avec les développements réalisés par la BNE pour son projet de la BDH.

Courant 2012, on procéda à l'intégration de différentes bibliothèques tout en continuant à améliorer le site lui-même (aussi bien dans son ergonomie que dans ses fonctionnalités). Il fut enfin rendu public en septembre 2012, à l'occasion de la XXIII^e Assemblée de l'ABINIA.

Quelques aspects techniques (dont les métadonnées)

L'une des caractéristiques de ce projet de la BDPI est qu'il s'appuie techniquement sur des standards ouverts et des logiciels libres, ce qui permet des économies non négligeables faites sur les licences et offre la possibilité de développer très librement les fonctionnalités qui apparaissent comme nécessaires au fil du temps. Sans entrer dans des considérations trop techniques, nous aborderons sommairement l'infrastructure technologique et les métadonnées utilisées dans ce projet.

La BDPI, une application basée sur les standards de J2EE, est programmée en Java et gérée par le serveur Apache Tomcat. Toutes les pages du moteur de recherche sont codées en HTML.

Les recherches sur le site se font grâce à la technologie Lucene/Solr qui se charge d'indexer les ressources fournies par les différentes bibliothèques participantes. Signalons que, comme c'est le cas pour d'autres sites similaires, ce ne sont pas les objets numériques mais les métadonnées sur ces objets (c'est-à-dire les informations bibliographiques les concernant) qui sont compilés dans la BDPI, l'URL menant à l'objet numérique conduisant donc au site de chacune des bibliothèques. La BDPI offre par conséquent cet avantage inestimable

6. <http://bdh.bne.es>

pour les bibliothèques participantes de rediriger les usagers de son portail vers chacun des « pourvoyeurs » de contenu.

Pour la « collecte » ou « récupération » de données (les métadonnées descriptives), c'est le protocole OAI-PMH qui est privilégié, mais le niveau de développement des bibliothèques participantes étant inégal, il a parfois été nécessaire d'avoir recours à des méthodes d'envoi différentes (courrier électronique ou autres systèmes de transfert de fichiers).

Une fois collectées toutes les métadonnées descriptives de toutes les sources, on utilise XSLT pour convertir ces métadonnées originales et les rendre accessibles à la plateforme de recherche SOLR.

Par ailleurs, un service de recherches externes (API), installé par Europeana il y a un certain temps, a été mis en place⁷, qui permet, depuis n'importe quelle interface, une recherche simultanée sur la BDPI. Ainsi, sur le site de la Biblioteca Digital Hispánica, par exemple, les deux services sont en présence, de sorte qu'une recherche lancée par un usager sur le site de la BDH s'exécute simultanément sur la BDPI et Europeana, pour une meilleure efficacité et rapidité d'information [voir illustration page précédente].

Il est intéressant de remarquer, alors que cette fonctionnalité n'a encore été implantée que sur le site de la BDH, que cette API se trouve être l'un des principaux accès à la BDPI (environ 14,72 % des accès se font via l'API). De plus, l'utilisateur qui accède à la BDPI par l'intermédiaire de l'API de la BDH reste presque deux fois plus longtemps et visite environ deux fois plus de pages que la moyenne des usagers du site. Il y a donc lieu de penser qu'il s'agit d'accès « de qualité », qui permettent à l'utilisateur de la BDH d'obtenir un précieux complément à sa recherche initiale d'information.

Indubitablement, dans la gestion d'un projet de ce type, la question des métadonnées est décisive. Pour le moment, le projet englobe un nombre

relativement réduit de bibliothèques ce qui a permis, une fois collectées les données (ou métadonnées) apportées par les bibliothèques participantes, de procéder à leur analyse détaillée : quantité et qualité des données, champs, format utilisé, identifiants et catégorisation du contenu structurel comme l'accès par matière. L'objectif est d'extraire le plus d'informations possibles pour une meilleure normalisation des données.

Suite à cette première analyse, les fonds ont été indexés selon douze critères permettant d'inclure un maximum d'informations et de répondre de manière plurielle à la recherche et à la récupération des données :

1. Auteur
2. Titre
3. Date
4. Matière(s)
5. Informations éditoriales
6. Identifiant
7. Type de document
8. ID
9. Description
10. Information afférente
11. Miniature
12. Institution

Les champs « date », « type de document » et « matière(s) » ont donné le plus de mal car, souvent très hétérogènes, ils sont particulièrement importants puisqu'on les utilise pour filtrer et limiter la consultation. Le cas du Brésil, par exemple, fut particulièrement complexe étant donné que les matières se trouvaient mentionnées en anglais et en portugais.

Quant au champ « miniature », il a fallu créer pour différentes institutions un procédé automatique permettant d'extraire la miniature de l'objet numérique original, lorsqu'on ne trouvait pas cette information parmi les métadonnées qui étaient fournies, alors même qu'il était fondamental de pouvoir la présenter à l'utilisateur sur l'interface.

Fonctionnalités

Si ce projet n'a pas eu jusqu'à maintenant l'appui de véritables financements, il a bénéficié du fait qu'il soit lié à la BDH, puisque son interface présente de solides fonctionnalités de

base, propices à une utilisation optimale du site.

Ces fonctionnalités permettent de faciliter la recherche et la récupération de l'information, mais aussi d'aider dans leur navigation les usagers qui entrent sans demande précise, curieux de se familiariser avec le contenu du site.

Ainsi, la recherche simple peut se faire à partir de chacun des champs, de façon rapide et intuitive. Cette recherche exclut en principe le contenu textuel des documents (l'OCR)⁸, afin d'éviter les « bruits » dans la réponse, mais il reste toujours possible d'inclure cette information, et ce de façon très simple.

La recherche avancée propose en outre de sélectionner plusieurs champs à la fois, et de restreindre par des filtres la consultation à une institution, un type de document ou une langue précise.

Les résultats apparaissent sous forme de liste abrégée (avec miniature, titre et auteur), classés par ordre de pertinence, mais il est aussi possible de les obtenir par titre, auteur ou date. On peut ensuite accéder directement au document (que l'on verra alors sur la page de l'institution qui le propose) ou à un aperçu détaillé qui fournit plus d'informations. Outre cette liste de résultats, possibilité est donnée, au moyen de filtres, de limiter les réponses par institution, matière, ou type de document.

Il existe une autre forme d'accès à l'information à partir d'une série de collections transversales basées sur des caractéristiques communes comme le type de document, le sujet, ou un intérêt ou une importance spécifique. Il s'agit là de donner à connaître d'une façon plus attrayante certains des trésors que recèle le site. Pour le moment, les collections qu'on peut explorer relèvent des domaines suivants : manuscrits, dessins, gravures et photographies, cartes, géographie et voyages, musique, littérature et études littéraires, presse et revues, contes et légendes.

7. Pour de plus amples informations, voir : www.bne.es/webdocs/RedesSociales/Blog/API.pdf

8. Reconnaissance optique de caractères (OCR correspondant aux initiales en anglais).

Utilisation et impact

Nous l'avons dit : un an après sa mise en place, il apparaît opportun de faire un premier bilan du portail pour en évaluer l'impact actuel mais aussi celui à venir tel que l'envisage la BDPI.

Au cours de cette première année, le portail de la BDPI a reçu 117 696 visites depuis 128 pays, 561 950 pages ayant été consultées. Les dix pays qui comptabilisent le plus d'accès sont l'Espagne, le Brésil, le Mexique, le Panama, la Colombie, l'Argentine, le Chili, les États-Unis, le Portugal et la France.

Si l'on s'attache aux entrées quotidiennes, la BDPI a maintenu au cours des premiers mois un nombre d'accès constant, autour de 400 visites par jour. Actuellement, le chiffre a baissé, il est plus proche de 250, même s'il faut signaler une recrudescence des visites juste après l'été.

Malgré ce chiffre assez bas, les conclusions tirées de l'observation des formes d'accès et de leur qualité poussent à l'optimisme. Ainsi, le temps moyen des connexions au cours de cette année était de 3,56 minutes. Si, bien sûr, on pourrait le souhaiter supérieur, le compte rendu du dernier trimestre d'Europeana, qui rapporte une moyenne de 15 000 à 20 000 visites journalières, fait état d'un temps d'accès sur le site européen de 2,18 minutes. En outre, au cours des deux derniers mois on relève sur la BDPI un temps moyen de 4,30 minutes (ce qui n'est pas négligeable, même si l'on est encore loin des 8,15 minutes de la BDH).

Autre élément positif : la diversité des sites affluents qui conduisent à la BDPI, qu'il s'agisse de réseaux sociaux, de blogs ou de sites consacrés à l'actualité. Incontestablement, les blogs culturels ont fait preuve d'une grande capacité à jouer un rôle de relais vers la BDPI. Il serait judicieux de se rapprocher de ce type de ressources afin de leur faire connaître et découvrir le projet.

Dans cette même optique, force est de constater l'utilité de l'API mentionnée plus haut. Depuis son implantation à la BDH, le trafic entre la BDH et la BDPI est resté constant, ce

qui prouve l'incontestable utilité de cet outil, pour la BDH elle-même qui élargit ses services en donnant accès à ses usagers à l'information de la BDPI (et d'Europeana), mais aussi pour la BDPI qui de la sorte conquiert de nouveaux usagers ; de nouveaux visiteurs qui, par ailleurs, font un usage accru du site, en termes d'intensivité et de « qualité » : face à la moyenne mentionnée de consultation, de 3,56 minutes, les usagers qui arrivent via l'API restent sur le site 7,18 minutes. Pour une moyenne de 4,8 pages par visite, eux en consultent 7,07. Quant au pourcentage d'abandon (le site est quitté dès la première page), il passe de 58,35 % à 38,3 %.

Autre chiffre significatif et potentiellement enrichissant : les visites recensées proviennent de 128 pays. On remarque que la présence de l'Espagne est particulièrement prédominante, sans doute due à une forte diffusion depuis les réseaux sociaux de la BNE (qui compte plus de 137 000 « followers » sur Facebook) et à l'implantation de l'API sur le site de la BDH.

L'Espagne, au cours de cette année, a ainsi représenté près de 40 % des visites. Il est intéressant d'observer que les 60 % restants sont répartis de façon très homogène sur un grand nombre de pays. Outre l'Espagne, les dix pays comptant le plus de « visiteurs » sont : le Brésil (21 %), le Mexique (10,5 %), le Panama (10 %), la Colombie (8,9 %), l'Argentine (7,9 %), le Chili (6,3 %), les États-Unis (5,2 %), le Portugal (3,7 %), la France (3,6 %) et le Pérou (3,4 %). On remarque que l'intérêt manifesté est presque égal dans différents pays et que cinq d'entre eux ne font pas partie de la BDPI.

Ces éléments, dont l'énumération peut sembler un peu fastidieuse, sont néanmoins très importants pour orienter le futur développement de la BDPI. À partir de l'analyse et de l'évaluation du comportement des usagers, il est possible d'élaborer les propositions adéquates qui permettront à la BDPI de réaliser un jour ce grand projet de compiler le contenu numérique culturel ibéro-américain.

Conclusion et prévisions

La conclusion essentielle que l'on peut tirer de cette première année de fonctionnement de la BDPI c'est, répétons-le, que le projet a fait preuve de solidité comme de flexibilité pour devenir un site de référence dans le domaine ibéro-américain. L'absence d'autres projets similaires place par ailleurs la BDPI dans une position privilégiée qui pourrait faire d'elle la première phase d'un grand projet compilateur du patrimoine numérique ibéro-américain.

En tout état de cause, dans une perspective plus concrète et plus proche d'une réalité immédiate, les prochaines étapes à mener doivent aller – et de fait, vont – dans le sens d'une augmentation du nombre de participants. Outre les bibliothèques membres de l'ABINIA, dont certaines ont déjà manifesté leur intérêt pour le portail et leur désir d'incorporation lors de la dernière réunion du groupe, l'éventuel ralliement de bibliothèques non nationales a été évoqué.

Bien entendu, ce serait un bond important qui devra être d'abord prudemment étudié puis mené à bien si et seulement si la démarche présentait les garanties nécessaires. Une adhésion croissante de participants est indispensable pour que le portail puisse amorcer ce saut qualitatif et quantitatif, et par conséquent souhaitable. Il n'en reste pas moins vrai que le cadre de l'ABINIA et de ses adhérents limite le volume des données et des institutions potentiellement participantes, rendant la gestion du projet beaucoup plus aisée.

De ce fait et dans une perspective d'avenir, en particulier si l'on souhaite faire de la BDPI un véritable outil de référence pour le patrimoine numérique ibéro-américain, il sera nécessaire de trouver un financement adéquat et durable. La question du financement apparaît comme un élément clé pour déterminer jusqu'où peut espérer se développer la BDPI. La portée de toutes les propositions qui seront faites pour l'avenir sera marquée par le budget qui aura ou non été obtenu. Il apparaît en tout cas que, grâce à l'utilisation de standards et de logiciels non propriétaires et à

l'expérience acquise dans des projets similaires, dont la BDH elle-même, gérer des projets importants avec des moyens limités reste concevable.

Toujours dans l'optique des améliorations à apporter, la BDPI, en plus de gagner de nouveaux participants, doit consacrer ses efforts à une diffusion plus massive. L'analyse des données d'utilisation du site nous révèle que son utilisation est plutôt bonne du point de vue de la «qualité». Néanmoins, le faible nombre de visites constitue un point sur lequel il faudra se concentrer à l'avenir.

Pour ce faire, il existe des solutions relativement simples qui passent essentiellement par une implication accrue de l'ABINIA et de ses membres dans la diffusion du projet (aussi bien pour les bibliothèques qui font déjà partie du réseau BDPI que pour les autres). Cette implication pourrait passer par une plus grande contribution de leur part à la diffusion de la BDPI via leurs voies de communication (les réseaux sociaux en étant le moyen le plus simple et économique) et par une implantation plus généralisée de l'API. De façon presque certaine, ces deux mesures si faciles à mettre en place augmenteraient significativement le nombre des accès au portail.

Il serait bon, également, d'entamer des dialogues qui révéleraient certainement des points de rencontre et des possibilités de collaboration et de diffusion avec des projets similaires comme lesdites Europeana et DPLA, mais aussi certainement beaucoup d'autres.

Enfin, les améliorations à apporter concernent le portail de la BDPI lui-même. Certes, il présente les fonctionnalités de base nécessaires à une bibliothèque en ligne et le comportement des usagers qui y accèdent est généralement satisfaisant, mais sans nul doute, avec des services élargis, on recenserait encore de meilleurs résultats. Ainsi, on pourrait envisager une alerte de documents, une meilleure mise en valeur des collections présentant un intérêt majeur, de propres canaux de diffusion, la possibilité de conserver des résultats de recherche, l'intégration de l'API d'Europeana, la rénovation de l'interface, la fixation des standards et des recommandations...

Comme nous l'avons dit, bien qu'il soit possible avec l'infrastructure actuelle de poursuivre le développement du site jusqu'à un certain point, pour que de vraies améliorations qualitatives et quantitatives voient le jour, un système de financement stable et suffisant sera nécessaire.

Étant donné toutes les réalisations menées à bien avec succès jusqu'ici, il devrait être possible de trouver ce financement et même un certain soutien politique, en vue de pérenniser le développement de la BDPI.

L'importance des relations historiques, culturelles ou littéraires entre les pays ibéro-américains doit se refléter dans la culture numérique des institutions mémorielles. La technologie nous permet de créer facilement des outils qui mettent à disposition dans ce nouvel espace les matériels patrimoniaux conservés, leur conférant ainsi une valeur différente de celle qu'ils ont en eux-mêmes.

En outre, un site de cette nature encourage la mise en place et la conduite de politiques de numérisation systématique des fonds, si utiles aux citoyens. Un point de rencontre qui sert aussi, presque inévitablement, d'élément de coordination et de normalisation, pour partager connaissances, expérience et technologie.

Pour toutes ces raisons, nous espérons voir la BDPI fêter encore de nombreux anniversaires ; puisse-t-elle se développer, dans les années à venir, à la mesure du potentiel qui est incontestablement le sien. ●

Octobre 2013