arbido

2013/4 Linked Open Data Big Data - Tout en réseau

Dubois Alain, Président de l'Association des archivistes suisses (VSA-AAS)

Inscrire les données du Service de la culture dans le Web pour mieux valoriser le patrimoine

Les questions relatives à la mise en réseau et à l'interopérabilité entre les systèmes d'information et de description, ainsi qu'à l'échange et/ou à la réutilisation des données constituent aujourd'hui un enjeu professionnel majeur pour les institutions de conservation du patrimoine.

Elles sont actuellement au cœur des réflexions menées par le Service de la culture du canton du Valais, dans le sillage de la mise en place d'une plateforme de pérennisation et d'archivage pour les collections numériques et numérisées des Archives de l'Etat du Valais, de la Médiathèque Valais et des Musées cantonaux conforme aux normes et standards internationaux en la matière. Elles aboutiront concrètement à la fin du premier semestre 2014 par la mise en ligne d'un outil de recherche fédérée dédié au patrimoine valaisan, l'ouverture d'un entrepôt OAIPMH (open data culturel) et l'alignement des notices descriptives selon les principes du Web de données (linked open data).

Le présent article retrace en trois parties les étapes successives d'un projet qui a débuté en 2011 et qui a pour ambition d'offrir à toute personne intéressée un accès aux ressources patrimoniales conservées par les Archives de l'Etat du Valais, la Médiathèque Valais et les Musées cantonaux à partir d'un portail unique.

1. L'outil de recherche fédérée du Service de la culture

L’origine du projet de recherche fédérée se situe en fait dans le sillage de l’ouverture de la plateforme de pérennisation et d’archivage en mai 2011. Il s’est agi, à cet instant, de mettre en réseau et de donner accès aux ressources numériques et numérisées disponibles sur cette plateforme, indépendamment des outils métier institutionnels usuels – scopeArchiv pour les Archives de l’Etat du Valais et le catalogue collectif RERO pour la Médiathèque Valais. Rapidement, toutefois, s’est imposée l’idée de réaliser un outil qui, au-delà des seules ressources numériques et numérisées, donne accès à toutes les ressources en lien avec le patrimoine conservées par le Service de la culture. C’est ainsi qu’est né le projet d’outil de recherche fédérée. Actuellement en voie de finalisation, ce projet s’est déroulé en deux phases:

– une phase-pilote (avril-décembre 2011), qui a permis de démontrer la faisabilité d’une mise en réseau de ressources disparates;

– une phase de déploiement (janvier 2012 juin 2014), qui permet de consolider les acquis de la phasepilote.

Sans revenir sur tous les éléments du projet, j’en mentionnerai ici les principaux.

Particularité de la structure de projet

Le projet est piloté par la cellule eculture du Service de la culture. Cette structure originale, qui regroupe des représen tants des Archives de l’Etat du Valais, de la Médiathèque Valais, des Musées cantonaux et de l’Encouragement des activités culturelles, a notamment pour objectif de mener des projets transversaux en lien avec les technologies numériques et relatives à la collecte, à la conservation, à la mise en valeur et à la promotion du patrimoine culturel valaisan et des ressources qui lui sont associées.

Particularité du projet

Dans sa philosophie, mais à une échelle nettement moindre, le projet de recherche fédérée se rapproche par ailleurs de la bibliothèque numérique européenne lancée en novembre 2008 (Europeana). Il s’agit, en effet, de mettre à la disposition du public intéressé par le patrimoine valaisan un outil qui fédère des ressources descriptives qui ne le sont pas encore (fonds d’archives, collections audiovisuelles, bibliothèque numérique et collections muséales). Sur ce point, le défi principal porte sur la mise en commun et l’alignement de ressources décrites selon des normes et standards qui diffèrent fortement.

Choix de l’outil informatique

La particularité du projet a eu une incidence significative sur le choix de la solution informatique. Au terme d'une procédure sur invitation, c'est l'entreprise française Antidot SAS qui a été retenue pour réaliser l'outil de recherche fédérée[1]. Ce choix s'est notamment justifié en raison des solides références dont disposait l'entreprise dans le domaine de la gestion de l'information, son expérience avérée dans la mise en œuvre d'outils de recherche fédérée et d'accès à l'information à travers de multiples portails et l'intégration des outils du web de données qu'elle proposait[L'entreprise Antidot a notamment réalisé le portail du Centre national de la recherche scientifique pour les sciences humaines et sociales (www.rechercheisidore.fr/ [consulté le 15 octobre 2013]).].

Choix du corpus

L'outil de recherche intègre actuellement 362 704 notices descriptives qui se répartissent comme suit: – 68 947 notices de fonds d'archives conservés aux Archives de l'Etat du Valais (Etat du Valais, communes, bourgeoisies et paroisses, ainsi que fonds privés) et progressivement mis à la disposition du public sur le portail scopeQuery http://scopequery.vs.ch/suchinfo.aspx [consulté le 15 octobre 2013].;

– 279 182 notices de la bibliographie valaisanne disponibles sur le portail valaisan de RERO;

– 14 225 notices des ressources audio visuelles conservées à la Médiathèque Valais – Martigny (photographies, enregistrements sonores et films) et mises à la disposition du public sur le portail «mémoire audiovisuelle du Valais»[2];

– 350 notices issues des deux catalogues «Collectionner au cœur des Alpes» publiés par le Musée d'art et le Musée d'histoire du Valais[3].

Livraison et traitement des données

L'un des principaux avantages du module Antidot Finder Suite, utilisé pour le traitement des données, réside dans le fait qu'il accepte en entrée des formats très variés, indépendamment de leur source. C'est ainsi que chaque office a pu déterminer son mode de livraison: les Archives de l'Etat du Valais exportent leurs unités de description en EAD (Encoded Archival Description), tandis que les notices de la bibliographie valaisanne sont livrées en Dublin Core et celles des collections audiovisuelles sont extraites de l'entrepôt OAI-PMH de la Médiathèque Valais – Martigny. Quant aux Musées cantonaux, ils fournissent leurs notices descriptives dans un fichier Excel.

Le module Antidot Finder Suite aligne ensuite les différents flux et les transforme en un format pivot dérivé du Dublin Core, avant d'appliquer sur le corpus ainsi constitué différents filtres d'indexation.

Editorialisation de l'outil de recherche fédérée

Le site dédié à la recherche fédérée est actuellement en cours de création C'est l'entreprise Conchita-Plus qui a été choisie pour créer l'interface de consultation (www.conchita-plus.ch/ [consulté le 15 octobre 2013]). A terme, il offrira au public un outil de re cherche dont les résultats pourront être affinés à l'aide de facettes (provenance institutionnelle, lieux, périodes historiques et personnes). Il mettra également en exergue un «document du mois» qui se fera l'écho de l'actualité d'un office (exposition, conférence, journée thématique ou encore nouvelle publication).

2. Vers un open data culturel

Dès les origines du projet, soucieux d'en assurer une diffusion aussi large que possible et dans le sillage des initiatives actuelles en la matière, le Service de la culture a décidé de rendre ses données interopérables et réutilisables. C'est ainsi que s'est très rapidement imposée l'idée de créer un portail qui rendent interopérables et moissonnables les données, dans la perspective de créer une plateforme d'open data culturel au niveau valaisan. Sur ce point, le Service de la culture a pu bénéficier de l'expérience de la Médiathèque Valais, qui rend disponibles depuis 2009 déjà ses collections audiovisuelles à travers un entrepôt OAI-PMHConsultable à l'adresse http://xml.memovs.ch/oai/, l'entrepôt OAI-PMH de la Médiathèque Valais rend disponibles sous forme de sets ses collections de photographies, d'enregistrements sonores et de films. Les données sont disponibles dans les formats suivants: oai_dc, qdc (Dublin Core qualifié), ead, marcxml et ese (Europeana) [consulté le 15 octobre 2013]. Ce modèle a donc été repris et adapté en fonction des besoins. C'est ainsi que l'entrepôt OAI-PMH du Service de la culture, actuellement en cours de validation auprès de l'*Open Archive Initiative*www.openarchives.org/ [consulté le 15 octobre 2013]. et mis à disposition au début de l'année 2014 sur le site Internet de l'Etat du Valais4, proposera ses données dans deux formats: le format oai_dc, constitutif de tout entrepôt OAIPMH, et le modèle de données d'Europeana (EDM).

Les sets seront ensuite rendus accessibles sous deux formes:

– par provenance institutionnelle (sets «Archives de l'Etat du Valais», «Bibliothèque numérique», «Mémoire audiovisuelle du Valais» et «Musées cantonaux»);

– par type de ressources (sets «Photographies», «Enregistrements sonores», «Films», «Monographies», «Articles» et «Fonds d'archives»).

3. L'intégration du linked open data

Rendre interopérables ses données à travers un entrepôt OAI-PMH constitue certes une première étape importante, mais elle reste néanmoins insuffisante au regard des possibilités qu'offre désormais le Web de données en matière d'échange et de réutilisation des données. C'est ainsi que l'alignement des notices descriptives selon le modèle de données d'Europeana permettra aux données du Service de la culture d'être présentées selon la structure du RDF (Ressource Description Framework).

Le référentiel des lieux, par exemple, sera aligné sur GeoNames, base de données géographique de plus de 6,5 millions de lieux existants www.geonames.org/ [consulté le 15 octobre 2013]. Des discussions sont par ailleurs en cours à l'interne pour aligner le référentiel des personnes avec le fichier d'autorités VIAF[5]. Par contre, l'abandon – temporaire toutefois – du référentiel des matières à ce stade du projet reporte l'alignement de ce dernier avec le thésaurus-matières pour l'indexation des archives locales proposé par le Service interministériel des Archives de France[6] ou la CDU abrégée[7].

Au final, les réflexions autour de la création d'un outil de recherche fédérée permettront non seulement d'offrir à toute personne intéressée un portail d'accès unique au patrimoine valaisan, qui pourrait du reste tout à fait être enrichi par des ressources en lien avec le Valais provenant d'autres institutions, mais surtout de l'inscrire dans le sillage des projets d'open data et de linked open data actuellement en cours en matière de valorisation du patrimoine. En somme d'inscrire pleinement les données du Service de la culture dans le Web pour les valoriser au mieux.

1 www.antidot.net/ [consulté le 15 octobre 2013].

2 http://archives.memovs.ch/ [consulté le 15 octobre 2013].

3 Ruedin, P., Collectionner au cœur des Alpes: le Musée d'art du Valais, Sion, Sion – Paris 2007; Elsig, P., Morand, M. C., Collectionner au cœur des Alpes: le Musée d'histoire du Valais, Sion, Sion – [Paris] 2013.

4 www.vs.ch/oai.

5 http://viaf.org/ [consulté le 15 octobre 2013].

6 http://data.culture.fr/thesaurus/page/ark:/67717/Matiere [consulté le 15 octobre 2013].

7 http://udcdata.info/ [consulté le 15 octobre 2013].

Alain Dubois

Licencié ès lettres et diplômé de l'École nationale des chartes, Alain Dubois est archiviste cantonal et directeur des Archives de l'État du Valais. Il a assumé la direction opérationnelle des Vallesiana (www.vallesiana.ch) jusqu'en 2020.

Depuis janvier 2021, il est également adjoint de la cheffe du Service de la culture du Canton du Valais.

Membre du Comité de l'Association des archivistes suisses depuis 2016, il la préside depuis 2020.

Portrait: © Olivier Maire, photo-genic.ch

Résumé

Deutsch

Die Überlegungen zur Onlinestellung und Interoperabilität zwischen verschiedenen Informationssystemen, ebenso wie der Austausch und/oder die Wiederverwendung von Daten stehen im Zentrum der Überlegungen der Dienststelle für Kultur des Kantons Wallis, die im Gefolge der Inbetriebnahme der Speicher- und Archivierungsplattform für die digitalen und digitalisierten Bestände des Staatsarchivs des Kantons Wallis, der Mediathek und der kantonalen Museen angestellt wurden. Am Ende des ersten Semesters 2014 sollen diese Überlegungen ganz konkret zur Onlinestellung einer gemeinsa-men Suche für das Walliser Kulturgut, zur Öffnung eines Datenlagers Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) und zum Abgleich der beschreibenden Daten nach den Grundsätzen des semantischen Webs führen.

Der vorliegende Artikel zeigt die wichtigsten Etappen des Projekts auf, das 2011 begonnen wurde und das zum Ziel hat, allen Interessierten den Zugang zu den kulturellen Ressourcen der Dienststelle für Kultur des Kantons Wallis über ein Guichet virtuel («One-Stop-Portal») zu ermöglichen.