Zeitschrift: Arbido

Herausgeber: Verein Schweizerischer Archivarinnen und Archivare; Bibliothek Information Schweiz

Band: - (2016)

Heft: 2: Ausgelagert, eingekauft, fremdbeschafft = Délocalisé, acheté, externalisé = Delocalizzato, comprato, esternalizzato

Artikel: Googlos - un partenariat public-privé pour la mise en ligne du patrimoine imprimé

Autor: Frey, Jeannette

DOI: https://doi.org/10.5169/seals-769413

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. Find out more

Download PDF: 14.07.2025

ETH-Bibliothek Zürich, E-Periodica, https://www.e-periodica.ch

Verwaltung digitaler Objekte in der Alma-Cloud, kann die BiG als Pilotpartnerin viel Einfluss nehmen und profitieren.

Nicht verschwiegen werden sollen jedoch auch einige negative Erfahrungen. Durch die direkte Verbindung in die Cloud besteht eine starke Abhängigkeit zu einem stabilen und leistungsfähigen Netz. Letzteres ist in der IT-Umgebung der BiG nicht in jedem Fall gegeben, was die Arbeit mit Alma gelegentlich ziemlich erschwert. Mit der Auslagerung in die Cloud hat das jedoch nur bedingt einen Zusammenhang. Zudem ist die Anpassung von Alma für den deutschsprachigen Raum noch nicht ganz abgeschlossen: Es fehlen noch einige Übersetzungen, und das Arbeiten mit der GND ist noch suboptimal.

Die Cloud ist letztlich nichts anderes als ein Server, dessen Standort bekannt ist.

Fazit

Bilanzierend kann festgehalten werden, dass die Erfahrungen mit dem Gang in die Cloud mehrheitlich positiv sind. Das hat jedoch mehr mit der Auslagerung des Betriebs an die Lieferantenfirma und damit einer deutlich schlankeren und effizienteren Struktur zu tun als mit der Cloud an sich. Denn die Cloud ist letztlich nichts anderes als ein Server, dessen Standort bekannt ist.

Der Erfolg eines Projekts steht und fällt im Kopf. Wird ein Projekt wie «BiG Portal» mit den notwendigen Sicherheiten und im Bewusstsein der Risiken (im übertragenen Sinn eben mit Regenschirm und Sonnenbrille) umgesetzt, ist man mental bereit für eine Auslagerung der Bibliotheksdaten in die Cloud, und steht man Anpassungen bisheriger gewohnter Betriebsabläufe an die neue Umgebung positiv entgegen, dann wird ein solches Projekt ein Erfolg.

Kontakt: Daniel.Kohler@gs-vbs.admin.ch, Rahel.BirriBlezon@gs-vbs.admin.ch

Bibliografische Angaben:
Kohler Daniel und Birri Blezon Rahel, «Die BiG geht in die Cloud – das Projekt BiG Portal», *Jahresbericht Bibliothek am Guisanplatz 2014*, 2014, S. 24–29, http://www.vbs.admin.ch/internet/vbs/de/home/departement/organisation/gensec/milit/veroffentlichungen.parsys.34473.downloadList.17007.DownloadFile.tmp/80091dgeschaeftsbericht big2014.pdf (Zugriff vom 15.3.2016)

ABSTRACT

Le réseau de bibliothèques Alexandria a remplacé, dans le cadre du projet BiG Portal et sous la direction de la Bibliothèque am Guisanplatz (BiG), son système de gestion des opérations de bibliothèque. Son choix s'est porté sur le système Ex Libris Alma après une phase d'évaluation intensive. L'exploitation et les données ont été délocalisées dans le cloud. Après cinq mois seulement d'implémentation, Alma a pu être introduit fin novembre 2013 dans le réseau Alexandria. Les expériences sont majoritairement positives : infrastructure plus légère et plus efficiente, système toujours à jour, gestion centrale de tous les types de documents (imprimés, numériques). Le plus grand défi est probablement posé par le fait qu'un système avec cloud nécessite un réseau stable et performant.

Googlos – un partenariat public-privé pour la mise en ligne du patrimoine imprimé

Jeannette Frey,
directrice BCU Lausanne

«Ici, la table de travail n'est plus chargée d'aucun livre. A leur place se dresse un écran, et à portée un téléphone. Là-bas, au loin, dans un édifice immense, sont tous les livres et tous les renseignements. De là on fait apparaître sur l'écran la page à lire pour connaître la réponse à la question posée par téléphone.»
Paul Otlet (1934), cité par Françoise Levie, *L'homme qui voulait classer le monde* (Bruxelles, 2006)

Entre autres missions légales, la BCU Lausanne collectionne le patrimoine écrit vaudois, par le dépôt légal ainsi que par les collections de la réserve précieuse, des manuscrits et des Valdensia. La mise à disposition de ces collections a toujours représenté un défi pour la BCU Lausanne, au vu des moyens personnels restreints. Depuis l'ouverture de l'Unithèque en 1982, la BCUL disposait cependant d'un atout majeur pour aborder le monde numérique : un catalogue en ligne complet de ses collections imprimées. Dès les années 1990, la BCU Lausanne s'est donc naturellement intéressée à raccrocher à son catalogue en ligne des documents numérisés. Cependant, le financement d'opérations de numérisation du patrimoine représentait – et représente encore un réel problème, les coûts de la numérisation de masse étant très importants.

En 2005, la BCU Lausanne a mené une réflexion stratégique inscrivant la « bibliothèque numérique » au nombre de ses axes de développement. La constitution d'une bibliothèque numérique vaudoise (livres, journaux, manuscrits) en était un des volets. Aucun financement extraordinaire n'étant alloué par l'autorité de tutelle, la BCU Lausanne s'est orientée vers un partenariat public-privé. Coïncidence, la réalisation de la vision de Paul Otlet citée en exergue se trouva être la vision an-

noncée par Google au lancement de son projet Google Books Search à la Foire du livre de Francfort en 2005. C'est ainsi que la BCU Lausanne en vint à établir des contacts avec Google en vue d'un partenariat dès 2006.

Un partenariat négocié

Si les intentions de Google et de la BCU Lausanne étaient claires dès le départ, les modalités d'interaction entre un acteur commercial global et une institution de l'Etat de Vaud ont nécessité d'intenses discussions durant l'année 2007. Ces négociations ont inclus les autorités de tutelle de la BCU Lausanne. Au final, un contrat de partenariat prévoyant la mise à disposition par la BCU Lausanne pour la numérisation par Google d'environ 106 000 volumes imprimés publiés avant 1870 a pu être signé en mai 2007. La BCUL devait mettre à disposition de Google les livres et les métadonnées (notices bibliographiques) : l'existence d'un catalogue en ligne était donc une condition sine qua non d'entrée en matière.

Les modalités d'interaction ont nécessité d'intenses discussions.

Née le 13 avril 1962 à Kirchberg (BE). Licence ès lettres (histoire ancienne, archéologie classique et égyptologie) à l'Université de Fribourg, puis rédactrice à la Rédaction LIMC (Lexicon Iconographicum Mythologiae Classicae) à Bâle. Responsable de la Photothèque du Musée national suisse à Zurich de 1992 à 1998 et premières expériences dans l'informatisation et la numérisation des fonds photographiques. Responsable des Archives fédérales des Monuments historiques à Berne de 1998 à 2992, en charge des projets d'informatisation et de numérisation de fonds photographiques. Responsable Information & Communication dans le secteur privé en 2002, tout en menant en parallèle des études d'informatique et de mathématiques à l'Université de Fribourg. En 2004, reprise de la responsabilité des périodiques et ressources électroniques à la BCU Lausanne. Directrice de la BCU Lausanne depuis mai 2008. Projets principaux en 2015: extension du bâtiment principal de la BCU Lausanne, l'Unithèque, sur le campus de Dorigny; RenouVaud, implémentation d'un nouveau SIGB pour l'ensemble des bibliothèques du réseau vaudois (plus de 100 bibliothèques).

Travailler avec Google

Est-il simple de travailler avec Google ? Oui et non. Tout d'abord, l'opération n'était pas entièrement dénuée de risques. Même une entreprise commerciale telle que Google peut faire faillite. Les autorités de tutelle de la BCU Lausanne ont donc exigé des sécurités pour ce cas de figure. Une opération de cette ampleur nécessite en outre la mise en place d'une logique de travail industrielle dont les processus ne ressemblent pas au quotidien d'une institution patrimoniale. Fort heureusement, la BCU Lausanne dispose, depuis la création de l'Unithèque, d'un centre de logistique permettant d'accueillir des camions de 28 tonnes, un prérequis pour assurer la logistique nécessaire. Quant au partenaire Google, il était déjà rôdé à l'exercice, et ses responsables ont su utilement seconder la BCU Lausanne dans la mise en place des processus et d'un workflow inédit.

Tout commence par une bonne planification des étapes avec Google et par le développement d'outils et produits informatiques particuliers. Ces développements ont été effectués par la BCU Lausanne en 2007. L'automne de la même année, les notices MARC des ouvrages entrant en ligne de compte pour le projet ont été extraits de la base de données RERO en format MARC-XML, et une base de données Xindice a été constituée. Les collections ont été découpées en lots et les signets d'accompagnement édités, les forces de travail nécessaires estimées et le personnel auxiliaire recruté.

Aucun livre n'a été perdu au cours de ces opérations.

Les résultats

Tous les ouvrages livrés par la BCUL ont-ils été scannés et mis en ligne par Google ? Pas tout à fait. Google se réservait le droit de ne pas mettre en ligne les scans des doublons. Dans le cas du projet Googlos, la BCU Lausanne a grandement bénéficié du fait d'être la première bibliothèque francophone à participer au projet Google Books Search. Le nombre de doublons préexistants était alors anecdotique, ce qui ne serait plus le cas aujourd'hui et engendrerait un processus de sélection que nous serions dans l'incapacité d'assurer.

Dans un partenariat de ce type, une institution patrimoniale espère obtenir des données numériques de la meilleure qualité possible. Un contrôle systématique de la qualité des images par la BCUL aurait cependant été bien trop lourd à gérer. Nous avons donc effectué des tests ponctuels, en fonction de certaines demandes. Google dispose de son côté de son propre système de contrôle des scans (« audit ») car tous les scans ne sont pas bons. Les erreurs les plus fréquentes concernent des pages tronquées (crop, 67%), des pages considérées comme vilaines (ugly, 15%), des problèmes dus au matériel (10%), des pages floues (blurry, 3%), plus rarement des pages fripées, biaisées, ou non alignées verticalement et horizontalement.

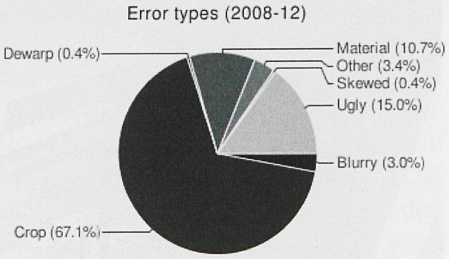

Image 1: Répartition des types d'erreurs.

La mise en oeuvre d'un workflow industriel permet de faire des constats intéressants quant à l'état général du catalogue et des collections.

Toutefois, le projet a pu être réalisé conformément au planning établi en 2007 et la numérisation finalisée en

Images 2, 3 et 4: Les livres sont préparés pour le transport afin d'être numérisés par Google.

Image 3

Image 4

2010. Plus de 110 000 volumes ont été envoyés par la BCU Lausanne, et quelque 107 000 scannés par Google. En 2012, les fichiers de 106 000 volumes avaient été traités par Google et étaient disponibles en ligne sur la plateforme de Google Books Search. Il a fallu pratiquement 2 ans à la BCU Lausanne pour rapatrier et archiver sur un système sécurisé les quelque 4 téraoctets de données que représentent les fichiers sources livrés par Google.

Bilan

Globalement, l'expérience a été très positive pour la BCU Lausanne, à preuve les nombreux témoignages reçus. Non seulement la collection est accessible par les internautes du monde entier, elle bénéficie aussi de la visibilité planétaire de la plateforme Google Books Search. En 2012, notre collection a généré plus de 120 millions d'accès par les internautes. Depuis 2013 et l'arrivée sur la plateforme des livres du projet de numérisation de la Bibliothèque municipale de Lyon, l'utilisation de la collection de la BCU Lausanne s'est stabilisée à environ 25 millions d'accès par année. Pour la recherche en histoire culturelle, la masse scannée constitue indubitablement une source directement exploitable, d'une taille ouvrant de nouvelles perspectives. Nous notons cependant un danger de donner trop de visibilité à des livres anciens (avant 1870!), parfois dépassés. Non scannées, les études plus récentes disparaissent de l'horizon du chercheur.

Accueillir des camions de 28 tonnes; un prérequis pour assurer la logistique nécessaire.

Durant le projet, la participation de la BCU Lausanne a suscité de vives réactions. Si le grand public y a toujours vu une avancée majeure dans l'accès à l'information écrite pour tous, partout et sans frais, les spécialistes de l'information et les chercheurs ont été partagés. Certains craignaient une situation de quasi-monopole d'une entreprise privée, d'autres dénonçaient une mainmise de la culture anglo-saxonne. D'autres encore dénonçaient une at-

teinte au droit d'auteur – ce qui n'a jamais été un argument dans le cas du projet de la BCU Lausanne, étant donné que le projet n'a concerné que la numérisation de collections hors droit.

Depuis, la BCU Lausanne a focalisé ses efforts sur la numérisation de la presse vaudoise, dont quelque 4 millions de pages sont déjà disponibles sur notre plateforme Scriptorium (http://scriptorium.bcu-lausanne.ch). Si l'utilisation n'atteint pas les sommets de notre collection Google Books Search, Scriptorium attire tout de même plus de 10 000 utilisateurs par jour en moyenne, générant quelque 3.5 millions d'accès par année.

Aucun livre n'a été perdu au cours de ces opérations.

En 2014, la BCU Lausanne a procédé à une nouvelle réflexion stratégique. Dans le nouveau plan directeur 2015–2020, la numérisation n'apparait plus guère que comme une note de bas de page. Comme nos acquisitions sont aujourd'hui à 2/3 numériques, et les processus de numérisation rôdés, nous nous concentrons principalement sur la maintenance des systèmes et des plateformes. Cependant, l'effort de numérisation reste conséquent et continuera pour un siècle au moins, par la numérisation de masse (externalisée) de la presse, la numérisation par objet (à l'interne, par l'intermédiaire d'un scanner à livre) des pièces isolées et des livres en format non standard, ou la numérisation spécialisée (effectuée par un photographe). Nous travaillons actuellement à la mise en place d'une nouvelle plateforme qui accueillera le processus du dépôt légal natif numérique, ainsi que les collections patrimoniales iconographiques et musicales.

Contact: jeannette.frey@bcu.unil.ch
Twitter: @JeannetteFrey

ABSTRACT

Die Bibliothèque Cantonale et Universitaire (BCU) Lausanne ist mit der Sammlung des schriftlichen Kulturgutes des Kantons Waadt beauftragt. Im Rahmen der strategischen Überlegungen 2005 wurde die Realisierung einer digitalen Bibliothek Teil der zukünftigen Entwicklungen. Da keine ausserordentlichen Mittel für die Finanzierung der Digitalisierung gesprochen werden konnten, machte sich die BCU Überlegungen zu den Möglichkeiten einer öffentlich-privaten Partnerschaft. Diese bot sich in der Zusammenarbeit mit Google und dem Projekt Google Books Search ab 2007. Der Artikel zeigt auf, wie die Zusammenarbeit mit Google zustande kam, wie die Digitalisierung umgesetzt wurde und welche Reaktionen diese Zusammenarbeit auslöste.

Sie wussten mehr! Vielen Dank!

«Offenes» Crowdsourcing im Bildarchiv der ETH-Bibliothek

Nicole Graf, Leiterin des Bildarchivs der ETH-Bibliothek

Nach dem erfolgreichen Crowdsourcing-Projekt mit Swissair-Pensionären (2009–2013) richtete das Bildarchiv der ETH-Bibliothek vor Kurzem auf seiner Bilddatenbank eine allgemein zugängliche Kommentarfunktion ein. Ein Artikel in der NZZ vom 18. Januar 2016 löste ein grosses Medienecho aus und rief viele Freiwillige auf den Plan, die mithelfen, Bilder zu identifizieren.

Das Bildarchiv der ETH-Bibliothek machte bereits zwischen 2009 und 2013 erste Erfahrungen mit einem Crowdsourcing-Projekt. Ehemalige Swissair-Mitarbeitende halfen ehrenamtlich mit, rund 40 000 Bilder des Fotoarchivs der Swissair online zu erschliessen[1]. In diesem Zusammenhang sprach die Autorin von Experten-Crowdsourcing, da gezielt eine geschlossene Gruppe von Experten angesprochen wurde. Der grosse Vorteil bei den Swissair-Pensionären war, dass sie gut organisiert und über diverse Kommunikationskanäle (eigene Zeitung, Versammlungen usw.) direkt ansprechbar sind.

Die Crowdsourcing-Gemeinde ist dankbar für Informationen.

Seit dem erfolgreichen Swissair-Projekt ist für uns klar, dass wir auch weiterhin auf das Wissen von vielen zurückgreifen wollen. Bilder ohne Titel können am effizientesten durch den sogenannten «Schwarm» identifiziert werden, das sechsköpfige Team des Bildarchivs hingegen hat ein begrenztes Wissen. Unter dem Motto «Wissen Sie mehr?» öffneten wir am 9. Dezember 2015 die Kommentarfunktion auf unserer Bilddatenbank Bildarchiv Online (http://ba.e-pics.ethz.ch) für jedermann, d.h., jeder Benutzende kann sämtliche Bilder via automatisierte E-Mail kommentieren.

1 Graf, Nicole: Experten erschliessen die Swissair-Bilder!, in: *arbido* 2 (2014), S. 37–39; Graf, Nicole: Crowdsourcing. Die Erschliessung des Fotoarchivs der Swissair im Bildarchiv der ETH-Bibliothek, in: Rundbrief Fotografie 23 (2016), No. 1 [N. F. 89], S. 24–32.