



## Construyendo puentes: de las bibliotecas Europeana a periódicos Europeana

**Susan K. Reilly**  
LIBER  
Países Bajos  
E-mail: susan.reilly[at]kb.nl

Session:

**119 – Users and portals: digital newspapers, usability, and genealogy – Newspapers Section with Genealogy and Local History Section**

### Compendio:

*Los estudios muestran que la facilidad de acceso y en particular la ventanilla inmediata (one-stop shop approach), son preferidos por los investigadores como una manera eficiente de acceder al contenido digital. Lo mismo que la facilidad de acceso, el contenido primordial de calidad garantizada es de principal importancia.*

*El proyecto Bibliotecas Europeana<sup>1</sup> aborda estos asuntos al seleccionar 5.1 millones de imágenes, libros, videos y estos y otros artículos directamente de 19 de las principales bibliotecas de investigación europea. La fuente de la información significa que se puede confiar en los metadatos y la calidad de la imagen, a la vez que la facilidad para acceder se garantiza por medio de una simple búsqueda de todos los objetos en la Biblioteca Europeana<sup>2</sup> y en los sitios de Europeana<sup>3</sup>.*

*Esta sociedad ha sentado las bases para una mayor colaboración e innovación en el futuro. Ya la infraestructura sostenible y la capacidad de búsqueda de textos completos por medio del proyecto de Bibliotecas Europeana se aplicarán a un nuevo contenido a través del proyecto de Periódicos Europeana<sup>4</sup>. Este proyecto tendrá disponible 29 millones de contenido de páginas e periódicos de toda Europa a través de las plataformas Biblioteca Europea y Europeana.*

---

<sup>1</sup> <http://www.europeana-libraries.eu>

<sup>2</sup> <http://www.theeuropeanlibrary.org>

<sup>3</sup> <http://www.europeana.eu>

<sup>4</sup> <http://www.europeana-newspapers.eu/>

## **Introducción:**

Los estudios muestran que la facilidad de acceso y en especial la ventanilla inmediata (one-stop shop approach), son preferidos por los investigadores como una manera eficiente de acceder al contenido digital. Igual que la facilidad de lograr esto, el contenido de gran calidad es de gran importancia.

Este documento presenta una parte del trabajo del proyecto de las Bibliotecas Europeana<sup>5</sup>, el cual permitirá que el texto completo de la biblioteca nacional y de investigación esté disponible en un solo portal. Ilustra los motivos y beneficios de la colaboración entre las organizaciones para lograr una visión común. También delinea los motivos y lo que se logrará en el proyecto de Periódicos Europeana, un proyecto que se beneficia del portal y la capacidad de búsqueda de textos desarrollados por medio de las Bibliotecas Europeana.

## **La piedra angular:**

El proyecto de Bibliotecas Europeana es un proyecto de dos años que empezó en enero de 2011. La idea para este proyecto surgió de la identificación de una necesidad de tener un punto común para las bibliotecas de investigación europeas, tanto nacionales como de universidades. Dichas bibliotecas habían trabajado juntas en el pasado para proveer un contenido temático para Europeana, el portal de herencia cultural Europeana por medio del proyecto de Viaje Europeana<sup>6</sup>. Dentro del contenido del proyecto en el tema de viajes y turismo se incluyen dos temas separados, uno para las bibliotecas nacionales y otro se estableció especialmente para aumentar el contenido de las otras bibliotecas en el proyecto. Dicho proyecto fue muy exitoso tanto en términos de proveer contenido digital de alta calidad a Europeana y en términos de establecer la colaboración entre bibliotecas nacionales y otras de investigación, pero surgió la pregunta sobre la sustentabilidad del uso de dos compiladores.

El proyecto Bibliotecas Europeana trata el asunto de la sustentabilidad al abrir el servicio de biblioteca nacional, la Biblioteca Europeana, a las bibliotecas de investigación. Utiliza este servicio para agregar una cantidad considerable de contenido valioso de las bibliotecas de investigación de Europa. Al final del proyecto, en diciembre de 2012, más de 5.1 millones de objetos, incluyendo 1.200 videos y películas, 850.000 imágenes y 4.3 millones de textos (libros, artículos periodísticos,

---

<sup>5</sup> <http://www.europeana-libraries.eu>

<sup>6</sup> <http://www.europeanatravel.eu/>

tesis, cartas) se habrán obtenido de las bibliotecas de investigación. Mucho de este contenido son textos y de especial valor para los investigadores. Para maximizar el potencial de este contenido, el proyecto se enfoca en desarrollar las capacidades de búsqueda de textos completos y un portal de búsqueda que provea las herramientas específicas para la investigación.

### **La sociedad:**

Al recabar el contenido de las bibliotecas de investigación y nacionales también se facilita el acopio de redes de bibliotecas europeas como CENL, CERL, y LIBER.

La Conferencia de Bibliotecas Nacionales Europeas (CENL) representa a 48 bibliotecas nacionales y actualmente es dueña de la Biblioteca Europea, la cual es la única biblioteca Europea de dominio. Hasta hoy sólo ha agregado contenido de las bibliotecas nacionales.

El enfoque del Consorcio de Bibliotecas de Investigación Europeas es mejorar el acceso y explotar la herencia europea impresa y cuenta con 33 miembros de bibliotecas de investigación y nacionales. CERL tiene el particular interés y experiencia para indexar y obtener metadatos.

LIBER, la asociación para las Bibliotecas de Investigación Europeas tiene más de 420 miembros (bibliotecas nacionales, universitarias y otras de investigación) de Europa y sus alrededores. 19 miembros de LIBER proveen el contenido para el proyecto y al final, el servicio desarrollado dentro de las Bibliotecas Europeana se extenderá a todos los miembros de LIBER.

Las Bibliotecas Europeana son la primera oportunidad para que estas tres organizaciones trabajen juntas con un firme propósito común, el cual es que sus instituciones miembro tengan un contenido que sea material de investigación valioso y todos deseen hacer accesible su contenido y fácil de usar para la comunidad dedicada a la investigación. Los resultados de la sustentabilidad del proyecto se asegurarán por medio del ejercicio de la disección de este punto en común y lo conviertan en proposiciones reales y convergentes y fortalezca la relación entre estas redes.

### **Definiendo el valor de la sociedad:**

El valor del proyecto reside no sólo en la creación de un servicio de una agrupación de servicios para bibliotecas, aun cuando éste sea un aspecto

importante, también reside en el potencial que ofrece para acercar el contenido de las bibliotecas a los investigadores en todo el mundo. Potencialmente extiende el alcance de las colecciones tanto de bibliotecas nacionales y de investigación más allá de las fronteras de sus comunidades y regiones de investigación establecidas. Explora la reputación colectiva de las bibliotecas como proveedores de confianza de información de calidad y de buenos metadatos. Es un hecho bien establecido que las bibliotecas están asociadas de manera muy positiva con los libros<sup>7</sup>. Al proveer el contenido completo de los textos de colecciones de libros digitalizados junto con otros contenido digitales tales como imágenes, videos y archivos de audio, sin mencionar aquellos de contenido académico como artículos y tesis, da como resultado que los investigadores pueden obtener resultados de investigación más provechosos. Al aumentar la visibilidad de dicho contenido de esta manera, las bibliotecas pueden aumentar el impacto de la importante inversión que hacen con la digitalización<sup>8</sup>.

Para los investigadores, el valor reside en poder tener acceso y examinar una importante cantidad de herencia cultural y de contenido de investigación relacionada en un solo lugar. Los estudios muestran que los investigadores están llevando a cabo búsquedas más complejas y también tienen menos tiempo para sus actividades de investigación<sup>9</sup>, de ahí el tener que hacer todo esto en un solo lugar es una proposición muy atractiva.

### **Diseño del portal:**

La mayor parte del contenido agregado por medio del proyecto de las Bibliotecas Europeana estará disponible a través del portal Europeana, pero tomando en cuenta el contenido que se está agregando, el portal de la Biblioteca Europeana fue rediseñado con el investigador de humanidades en mente. Dicha proposición es particularmente relevante

---

<sup>7</sup> De Rosa, Cathy et al. 2005. Perceptions of Libraries and Information Resources: A Report to the OCLC Membership. Dublin, Ohio: OCLC. <<http://www.oclc.org/reports/2005perceptions.htm>>.

<sup>8</sup> European Commission, Maurice Lévy, Elisabeth Niggemann, and Jacques de Decker. The New Renaissance. Brussels: European Commission, 2011. [http://ec.europa.eu/information\\_society/activities/digital\\_libraries/doc/reflection\\_group/final\\_report\\_%20cds.pdf](http://ec.europa.eu/information_society/activities/digital_libraries/doc/reflection_group/final_report_%20cds.pdf)

<sup>9</sup> Bulger, Monica, et al., Reinventing Research? Information Practices in the Humanities. London: Research Information Network, 2011.

para la comunidad de investigadores dedicados a las humanidades, para quienes la definición de datos de investigación es compleja:

*“La comunidad de las ciencias relacionadas con las humanidades necesita una importante cantidad de recursos digitales lo mismo que herramientas comunes, servicios y depósitos si han de avanzar más allá del concepto ‘boutique de proyectos’ a una base sólida de teoría y métodos”*<sup>10</sup>

Una serie de talleres y de pruebas por parte de los consumidores, lo mismo que la investigación efectuada, ha dado como resultado el diseño de un nuevo portal que incluye:

- la habilidad para buscar contenidos completos
- la oportunidad de inspeccionar el registro de metadatos de objetos individuales con la meta de lograr el acceso a grandes bases de datos con el propósito de investigarlos.
- los cada vez más utilizados CERIF como encabezados. Esto hace posible la investigación en una amplia gama de corpus de referencia unidos por un tema común o período de tiempo.
- Desarrollo de colecciones Pan-Europeas en cooperación con las bibliotecas nacionales, de investigación y universitarias de Europa. Esto es una extensión de las exhibiciones virtuales actuales<sup>11</sup>, las cuales muestran el contenido de una amplia gama de fuentes en toda Europa.
- Líneas de tiempo que muestran la frecuencia con que un término en particular es buscado a través de los siglos.
- APIs permitirá que el contenido de la base de datos se analice y se muestre en contextos fuera del portal de la Biblioteca Europea. Esto significa que el investigador puede llevar su contenido a su propia investigación y explorará nuevas formas de analizarlo.
- La exportación directa de registros a los servicios de administración de referencia popular tales como Mendeley y Zotero.

Una vez lanzado, el portal estará constantemente re-desarrollándose en línea con otras investigaciones emergentes en el área de ciencias humanas y digitales. Se harán estudios subsecuentes para saber de qué

---

<sup>10</sup> Christine L. Borgman. "The Digital Future is Now: A Call to Action for the Humanities" Digital Humanities Quarterly 4.1 (2010). Available at: <http://works.bepress.com/borgman/233>

<sup>11</sup> [www.theeuropeanlibrary.org/exhibition](http://www.theeuropeanlibrary.org/exhibition)

manera el investigador puede explotar y utilizar este contenido y estrategia de contenido para desarrollarlo en línea.

### **Edificando puentes:**

Además de los servicios de agrupamiento, el portal y la búsqueda de textos completos desarrollados en el proyecto de las Bibliotecas Europeana se usarán para exponer un nuevo tipo de contenido: periódicos. Desarrollos recientes en el OCR hechos a través del proyecto IMPACT<sup>12</sup> se aplicarán al contenido periodístico de 12 bibliotecas de investigación nacionales en todo Europa. 18 millones de páginas de periódicos serán filtradas y publicadas en el portal de Europeana y la Biblioteca Europea. El proyecto de Periódicos Europeana saldrá para cubrir los desafíos específicos que permiten que el texto completo de los periódicos viejos esté disponible. Hará uso de métodos de depuración para la segmentación de OCR, OLR/segmentación de artículo y reconocimiento (NER) de la página de la clase.

Mucho de lo que se ha desarrollado y aprendido a través de las Bibliotecas Europeana se aplicará a los Periódicos Europeana:

#### 1. Agrupación:

Se pueden identificar cuatro tipos de colecciones de periódicos digitales existentes:

- a) Imágenes sólo con metadatos estructurales
- b) Imágenes con metadatos estructurales y texto completo para investigación (OCR)
- c) Imágenes con metadatos estructurales, reconocimiento de artículos (OLR/rastreo de artículos) y OCR
- d) Imágenes con metadatos estructurales, OLR, OCR y enriquecimiento semántico. Todos los datos disponibles serán recabados por la Biblioteca Europea. La información será transformada a EDM, información modelo de Europeana y distribuída para Europeana.

#### 2. Estandarización de los metadatos.

Actualmente se encuentra en uso una gran variedad de formatos de metadatos. Para mejorar el acceso al contenido digital, se deberán adoptar las

---

<sup>12</sup> <http://www.impact-project.eu/>

normas comunes. Todos los formatos de metadatos existentes serán identificados y se darán las mejores soluciones a la comunidad.

### 3. Capacidad para mejor visualización.

Lograr que los periódicos sean fáciles de investigar y presentarlos de manera atractiva en línea es algo que actualmente es un desafío. El proyecto Periódicos en línea tendrá como modelo el trabajo hecho por las Bibliotecas Europeana y de ahí partirá para especificar la búsqueda apropiada y los requisitos de presentación, que puede usar Europeana.

El texto de periódicos del pasado, incluso desde el siglo XVIII, estará disponibles en línea. Los usuarios podrán ver estos periódicos en contexto, junto con imágenes de arte, fotografías, tesis relevantes, libros y artículos. Al presentar el contenido del periódico de esta manera se podrán hacer nuevas conexiones y se abrirán puertas para nuevos tipos de investigación y colaboración.

Existen otros beneficios para el trabajo de este proyecto como los procedimientos con los cuales el contenido del periódico será mejorado y que incluye el OCR, OLR/búsqueda de artículos, NER y reconocimiento de la página. Para cada uno de estas tareas técnicas lo mejor que se pueda hacer será identificado y publicado. Esto será de gran beneficio para tener una mayor red de bibliotecas con las redes CERL, CENL y LIBER. Esto ayudará a reducir el costo de los proyectos de digitalización de los periódicos y aumentará el acceso a las colecciones de periódicos digitales ahora y en el futuro.

### **Conclusión:**

Las Bibliotecas Europeana fueron un trabajo de grupo que abordó la necesidad de que se busque un recopilador para las bibliotecas europeas. Al hacerlo se logró unificar las redes claves de bibliotecas. Al trabajar hacia una visión en común, las redes han creado una fuente que tendría un gran valor potencial para la comunidad dedicada a la investigación. También creó las condiciones para que las bibliotecas nacionales y de investigación trabajen juntas de manera más fluida, enfocándose en la visión de conectar el contenido y al investigador. El contenido que se tiene en las bibliotecas europeas es rico y variado. Esto es particularmente cierto para las compañías de periódicos. Contar con estos compañías en línea en un momento en que la tecnología se está especializando y desarrollando para mostrar el texto completo de manera significativa para crear una oportunidad para los investigadores de que interactúen y se conecten con el vasto y rico acervo cultural europeo.

También es la incorporación de organizaciones, sistemas e instituciones que colaboran y pueden producir resultados innovadores, mejorando la eficiencia y dando como resultado una mayor accesibilidad. Tal logro tendrá resultados a largo plazo, no solamente para las bibliotecas o para la accesibilidad a la herencia cultural europea, sino para cada país en el mundo que cuenta con una gran cantidad de material cultural impreso.

## **Referencias:**

Bulger, Monica, et al., *Reinventing Research? Information Practices in the Humanities*. London: Research Information Network, 2011.

Christine L. Borgman. "The Digital Future is Now: A Call to Action for the Humanities" *Digital Humanities Quarterly* 4.1 (2010). Available at: <http://works.bepress.com/borgman/233>

De Rosa, Cathy et al. 2005. *Perceptions of Libraries and Information Resources: A Report to the OCLC Membership*. Dublin, Ohio: OCLC. <<http://www.oclc.org/reports/2005perceptions.htm>>.

European Commission, Maurice Lévy, Elisabeth Niggemann, and Jacques de Decker. *The New Renaissance*. Brussels: European Commission, 2011.  
[http://ec.europa.eu/information\\_society/activities/digital\\_libraries/doc/reflection\\_group/final\\_report\\_%20cds.pdf](http://ec.europa.eu/information_society/activities/digital_libraries/doc/reflection_group/final_report_%20cds.pdf)

Europeana (2012) Retrieved from <http://www.europeana.eu> on 19<sup>th</sup> April 2012.

Europeana Libraries (2012), Retrieved from <http://www.europeana-libraries.eu>, on 9<sup>th</sup> March 2012.

Europeana Travel (2012), Retrieved from <http://www.europeanatravel.eu/> on 19<sup>th</sup> April 2012.

IMPACT, Retrieved from <http://www.impact-project.eu/> on 19<sup>th</sup> April 2012.

The European Library(2012), Retrieved from <http://www.theeuropeanlibrary.org> on 9<sup>th</sup> March 2012