
GREENSTONE: ACTUAL USE IN ARGENTINA

GREENSTONE: ACTUAL USE IN ARGENTINA

GREENSTONE: USO ATUAL NA ARGENTINA

¹Marcela Fushimi, ¹Mariana Pichinini, ²Ana María Sanllorenti, ²Martín Williman
¹Universidad Nacional de La Plata, ² Universidad de Buenos Aires

Correspondência

¹Marcela Fushimi
Universidad Nacional de La Plata
Buenos Aires, Argentina
Email: mfushimi@fahce.unlp.edu.ar
ORCID: <https://orcid.org/0000-0002-7266-8216>

Submetido em: 05/03/2018

Aceito em: 25/03/2018

Publicado em: 23/04/2018



JITA: HS. Repositories

RESUMEN: Este trabajo presenta las características e historia del software para crear y gestionar bibliotecas digitales Greenstone, desarrollado inicialmente por la Universidad de Waikato, Nueva Zelandia en 1997. En primer lugar, se describe la comunidad de usuarios a nivel global, focalizando en el uso actual que el software ha alcanzado en Argentina: la cantidad de implementaciones disponibles, su evolución, el tipo, tamaño y variedad de los desarrollos existentes, así como su aplicación a la gestión de repositorios digitales de ciencia y tecnología en el ámbito de las instituciones científicas, tecnológicas y de educación superior en Argentina. En segundo lugar, se detallan las acciones llevadas a cabo a partir de la creación del Centro Nacional de Promoción de Greenstone en Argentina en 2009. Datos recabados en encuestas realizadas permitieron observar que las razones predominantes para elegir esta plataforma fueron, entre otras, su facilidad de instalación y configuración, su bajo nivel de requerimiento tecnológico, la generalizada escasez de recursos humanos dedicados a esta actividad, y la complejidad que presentaban los softwares alternativos existentes en ese momento. A lo largo de estos 8 años, tanto los repositorios digitales como los sistemas que los soportan evolucionaron drásticamente, modificando el escenario actual. Paralelamente, en 2016 la nueva versión mayor de Greenstone implementó una reingeniería completa del software para su adaptación a las tecnologías en uso: XML, XSLT, Web-services y Java. Como consecuencia de esto, la comunidad de desarrollo local se plantea nuevos desafíos para la migración de las bibliotecas digitales y repositorios implementados con versiones anteriores.

PALABRAS CLAVE: Repositorios institucionales – Argentina. Biblioteca digital. Biblioteca universitaria. Software de código abierto. Desarrollo de la capacidad.

ABSTRACT: This paper presents the characteristics and history of the software Greenstone to create and manage digital libraries, initially developed by the University of Waikato, New Zealand in 1997. First, the user community is described at a global level, focusing on the current use that the software has achieved in Argentina: the number of implementations available, their evolution, the type, size and variety of the existing developments, as well as their application to the management of digital repositories of science and technology in the field of scientific, technological and higher education institutions in Argentina. Secondly, the actions carried out after the creation of the National Center for the Promotion of Greenstone in Argentina in 2009 are detailed. Data collected from surveys made it possible to observe that the predominant reasons for choosing this platform were, among others, its ease of installation and configuration, its low level of technological requirements, the general scarcity of human resources dedicated to this activity, and the complexity of the alternative software existing at that time. Throughout these 8 years, both the digital repositories and the systems that support them evolved drastically, modifying the current scenario. In parallel, in 2016 the new major version of Greenstone implemented a complete re-engineering of the software to adapt it to the technologies in use: XML, XSLT, Web-services and Java. As a consequence of this, the local development community poses new challenges for the migration of digital libraries and repositories implemented with previous versions.

KEYWORDS: Institutional repositories – Argentina. Digital libraries. Academic Libraries. Open source software. Capacity building.

RESUMO: Este artigo apresenta as características e a história do software para criar e gerenciar bibliotecas digitais Greenstone, desenvolvidas inicialmente pela Universidade de Waikato, Nova Zelândia, em 1997. Primeiro, a comunidade de usuários é descrita a nível global, com foco no uso atual que o software conseguiu na Argentina: o número de implementações disponíveis, sua evolução, o tipo, tamanho e variedade dos desenvolvimentos existentes, bem como a sua aplicação à gestão de repositórios digitais de ciência e tecnologia no campo das instituições científicas, tecnológicas e de ensino superior na Argentina. Em segundo lugar, são detalhadas as ações realizadas após a criação do Centro Nacional para a Promoção da Greenstone na Argentina em 2009. Os dados coletados das pesquisas permitiram observar que os motivos predominantes para a escolha desta plataforma foram, entre outros, a facilidade de instalação e configuração, seu baixo nível de requisitos tecnológicos, a escassez geral de recursos humanos dedicados a essa atividade e a complexidade do software alternativo existente naquele momento. Ao longo destes 8 anos, tanto os repositórios digitais quanto os sistemas que os suportam evoluíram drasticamente, modificando o cenário atual. Paralelamente, em 2016, a nova versão principal da Greenstone implementou uma reengenharia completa do software para adaptá-la às tecnologias em uso: XML, XSLT, serviços Web e Java. Como consequência disso, a comunidade de desenvolvimento local apresenta novos desafios para a migração de bibliotecas digitais e repositórios implementados com versões anteriores.

PALAVRAS-CHAVE: Repositórios institucionais – Argentina. Biblioteca digital. Biblioteca universitaria. Software livre. Desenvolvimento de capacidade.

1 INTRODUCCIÓN Y METODOLOGÍA

Este trabajo introduce las características e historia del software para crear y gestionar bibliotecas digitales Greenstone, desarrollado por la Universidad de Waikato en Nueva Zelanda y distribuido bajo la Licencia Pública General (GNU) desde 1997.

La primera parte presenta un relato histórico que reconstruye a partir de fuentes documentales y testimonios de los protagonistas, la forma en que el software se introdujo en los países de América Latina y especialmente en Argentina, gracias al apoyo de la Unesco. Luego se detallan las acciones llevadas a cabo a partir de la creación en 2009 del Centro Nacional de Promoción de Greenstone (CNG) en Argentina.

En la segunda parte del artículo, en base a un relevamiento ad-hoc realizado a partir de la observación y análisis de los sitios web hechos con Greenstone en Argentina, se describe el uso actual que el software ha alcanzado: la cantidad de implementaciones existentes, su evolución, el tipo, tamaño y variedad de las colecciones desarrolladas, así como su aplicación a la gestión de repositorios digitales de ciencia y tecnología en el ámbito de las instituciones científicas, tecnológicas y de educación superior en Argentina.

En la tercera parte se exponen las características de la versión mayor 3 lanzada en 2016, que implicó una reingeniería completa del software para incorporar nuevas tecnologías de la web, y a continuación se describen dos casos de bibliotecas digitales implementadas con la nueva versión.

Finalmente, se realizan consideraciones generales a partir de los usos observados y las posibilidades ofrecidas por la versión 3, a la luz de los cambios y avances ocurridos tanto en los repositorios digitales como en los sistemas que los soportan.

2 CARACTERÍSTICAS E HISTORIA DE GREENSTONE

Greenstone es un conjunto de programas de software diseñado para crear y distribuir colecciones digitales, que permite organizar y publicar la información a través de Internet o en formato CD-ROM. Se originó en el Proyecto Biblioteca Digital de Nueva Zelanda con sede en la Universidad de Waikato, y fue desarrollado en colaboración con la UNESCO y la ONG de Información para el Desarrollo Humano con sede en Amberes, Bélgica. Es un software libre, de código abierto, multilingüe y multiplataforma, distribuido a partir de 1997 conforme a los términos de la Licencia Pública General GNU (GREENSTONE DIGITAL LIBRARY SOFTWARE).

El objetivo inicial de los desarrolladores fue ofrecer una herramienta sencilla y completa que permitiera a instituciones de todo tipo y tamaño organizar sus colecciones digitales y ofrecerlas a sus lectores. Todo ello sin la necesidad de contar con una potente

infraestructura y equipamiento -bastaba inicialmente con una PC de escritorio- ni tampoco conocimientos informáticos expertos para poder instalar y configurar el programa. Cualquier persona con nociones básicas de computación y ganas de explorar era capaz de crear y gestionar una biblioteca digital en unos pocos pasos. Esto generó que el programa se difundiera ampliamente.

Greenstone acepta documentos en una amplia gama de formatos tanto propietarios como estándares, y es compatible con numerosas normas de intercambio de documentos y metadatos, incluyendo el cumplimiento del protocolo OAI-PMH¹ y las normas Z39.50 de recuperación de información. Convierte rápidamente bases de datos bibliográficas en bibliotecas digitales, incluyendo los textos completos de los documentos de referencia si están disponibles.

Una de sus principales premisas ha sido asegurar que el software pone mínimas exigencias de recursos al sistema y es de fácil instalación, así como mantener la compatibilidad con versiones anteriores. La flexibilidad, robustez, facilidad de uso y libre disponibilidad lo convirtieron en un recurso especialmente útil para los países en desarrollo.

El núcleo del sistema (hasta la versión 2 inclusive) está programado en C++ y el funcionamiento general utiliza scripts desarrollados en Perl. Posee un motor de búsqueda propio (MGPP), aunque también trabaja con Lucene de Apache. Cuenta con tres interfaces: una interfaz gráfica para la administración, configuración y actualización de la biblioteca digital, una interfaz por línea de comandos para el mantenimiento interno y la interfaz web para publicar la biblioteca digital en Internet.

Paralelamente al mantenimiento y actualización de la versión 2, en 2002 el equipo de la Universidad de Waikato inició el desarrollo de la versión 3, que incluyó una reingeniería completa del software con migración de tecnologías y nuevas funcionalidades. Entre 2000 y 2006 la Universidad de Waikato trabajó en desarrollar documentación, dictar talleres de capacitación en varios países, generar tutoriales y ejemplos de uso para promover la adopción de la nueva versión (GREENSTONE WIKI).

Como en todos los proyectos de código abierto, la base de usuarios de Greenstone es desconocida, ya que se distribuye a través del sitio SourceForge². Sin embargo, es posible dimensionarlo a través de algunos datos puntuales, como ser:

¹Open Archives Initiative Protocol for Metadata Harvesting

²<https://sourceforge.net/projects/greenstone/?source=directory>

- La distribución desde SourceForge inició en noviembre de 2000.
- El promedio de descargas por mes desde entonces ha sido de 5000, mientras que el total de descargas registradas hasta junio de 2015 era cercana al millón.
- El software se descargó de 170 países diferentes y la lista de usuarios registrados es de 770.

En relación a su uso en nuestra región, 2005 marcó un hito importante a partir de la realización del “*Curso Mercosur para la construcción de bibliotecas digitales*” (UNESCO, 2005), que contó con el apoyo de la Oficina Regional de Ciencia de la UNESCO para América Latina y el Caribe, la Biblioteca Nacional de Uruguay, y tuvo el auspicio de la Universidad de la República Oriental del Uruguay. El curso se realizó en Montevideo, la Unesco financió los honorarios y estadías de los profesores, y cada país seleccionó a los participantes, que asistieron becados por sus instituciones.

Los docentes de ese curso fueron Jesús Tramullas, profesor de la Universidad de Zaragoza, España; Ana María Sanllorenti, en aquel momento de la Biblioteca Nacional Argentina; Ximena Cruzat, de la Biblioteca Nacional de Chile; Ana Pavani, de la Universidad PUC de Brasil, y los profesores Fernando Da Rosa y Rodolfo Pilas, de la Universidad de la República de Uruguay y del Grupo de Usuarios Linux de Uruguay respectivamente. Gracias a ese curso, el software comenzó a conocerse en los países de América Latina, ya que asistieron al mismo 62 bibliotecarios y profesionales de Argentina, Brasil, Chile, Colombia, Perú, Paraguay y Venezuela.

En diciembre de 2008 Greenstone fue premiado por la Fundación Andrew Mellon de EEUU debido a su aporte al ámbito de la educación, la cultura y la promoción humana. Según el representante de Unesco para América Latina y del Caribe:

“Este premio ha permitido a Greenstone, por primera vez, tener sus propios recursos discrecionales para ayudar a poner en marcha iniciativas en los países en desarrollo y para complementar el apoyo de los donantes internacionales” (CYRANEK, G., 2010, p.9).

La Universidad de Waikato recibió el premio y destinó parte del dinero para la difusión y el fortalecimiento de las comunidades de usuarios y desarrolladores de Greenstone, particularmente en países en vías de desarrollo en Latinoamérica. Tomando como base la lista de usuarios en español, se publicó una convocatoria para la creación de centros nacionales que se encargarían de la promoción local del software.

Dichos centros debían conformarse por instituciones sin fines de lucro que tendrían como objetivos: promover el software en su país o región asociada, realizar un relevamiento para conocer el uso en su país, su ámbito de aplicación, su grado de aceptación y los posibles requerimientos, crear un directorio de especialistas locales, definir un calendario de talleres de capacitación, colaborar con el proceso de traducción de la documentación al español,

desarrollar un sitio de acceso libre para alojar al Centro Nacional y fomentar la colaboración regional.

En base a esta convocatoria, se crearon 4 Centros Nacionales en la región: dos en Chile (Universidad Cardenal Silva Henríquez UCSH y Fundación para la Innovación Agraria (FIA), uno en Cuba (Centro de Cibernética Aplicada a la Medicina CECAM) y otro en Argentina.

3 CENTRO NACIONAL DE PROMOCIÓN DE GREENSTONE ARGENTINA (CNG)

En Argentina, la iniciativa conjunta de la [Biblioteca Prof. Guillermo Obiols](#) de la Facultad de Humanidades y Ciencias de la Educación de la Universidad Nacional de La Plata y la [Biblioteca Central Dr. Luis Federico Leloir](#) de la Facultad de Ciencias Exactas y Naturales de la Universidad de Buenos Aires, constituyeron el Centro Nacional de Promoción de Greenstone Argentina (CNG) en diciembre de 2009.

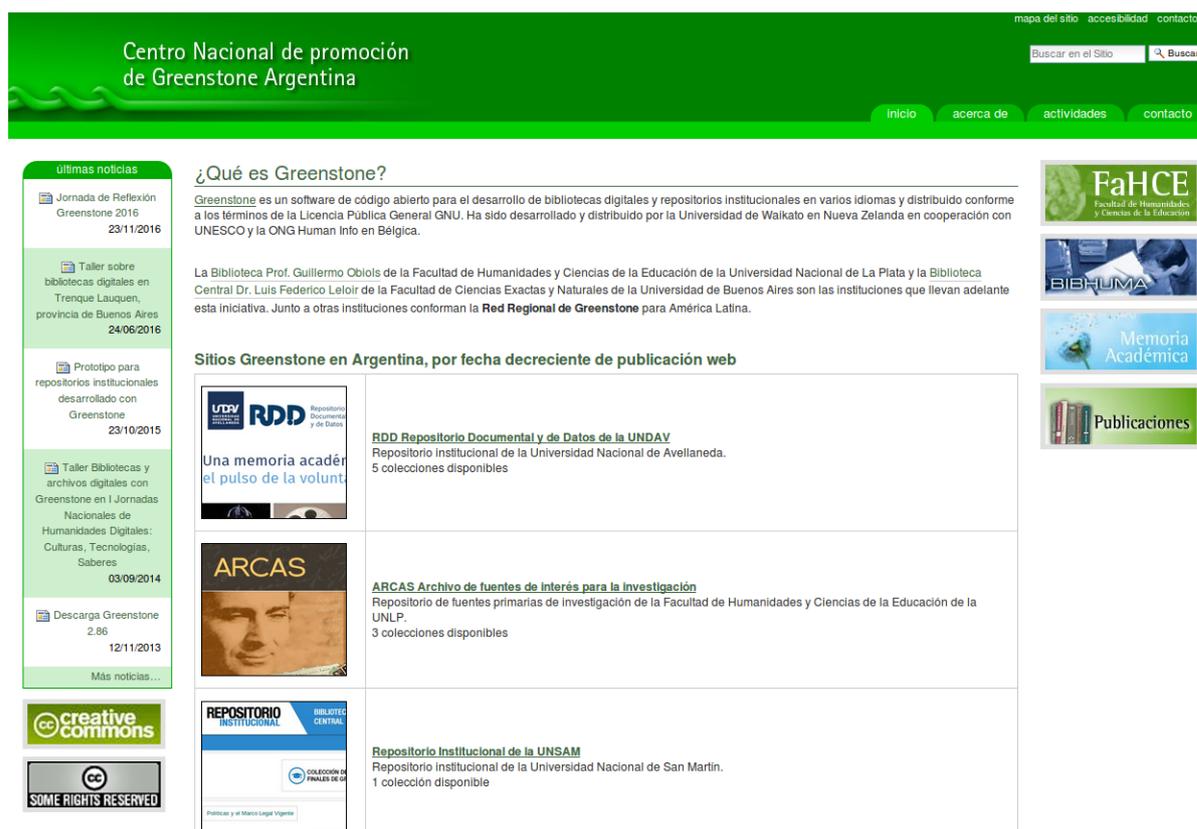


Figura 1. Página de inicio del CNG

Fuente: Recuperado de <http://cng.fahce.unlp.edu.ar/> el 25-2-2018.

Entre las actividades desarrolladas en el período 2010-2016 merecen destacarse las siguientes: la realización de dos reuniones nacionales de usuarios de software (2009 y 2011), tres encuestas sobre el desarrollo de bibliotecas digitales y el uso de Greenstone en Argentina

(2009, 2011 y 2016), una cuantiosa cantidad de cursos, talleres y charlas de difusión dictados, la creación y actualización permanente del sitio web del Centro Nacional (<http://cng.fahce.unlp.edu.ar>), el desarrollo de un prototipo para la creación de repositorios institucionales, y la asesoría y soporte técnico especializado a instituciones y usuarios que lo requirieron.

El CNG dictó 8 cursos y talleres de capacitación de tipo introductorio, intermedio y avanzado entre 2009 y 2014, todos gratuitos y con el auspicio de instituciones anfitrionas que cedieron el espacio. En 2016 organizó asimismo una Jornada de Reflexión para las instituciones que están usando actualmente el software, donde se debatió sobre la situación actual de cada una y las perspectivas futuras.

En el marco de la participación en un Proyecto denominado “Investigación y Desarrollo en Repositorios Institucionales de las Universidades Nacionales de la Región Bonaerense” (PICTO-CIN 149 2011/2012), desde el CNG se desarrolló un prototipo para la creación de repositorios institucionales basados en Greenstone (PICHININI, M., 2012), como una forma de aportar recursos a las instituciones participantes. Teniendo en cuenta que Greenstone es un software genérico para bibliotecas digitales, el objetivo del prototipo fue reducir el costo de adaptación y configuración del software para la gestión de un repositorio institucional tipo, con la provisión de tres colecciones básicas ya desarrolladas (tesis, artículos y ponencias). Las tres colecciones incluyen un esquema de metadatos específico para cada tipo documental y el esquema de metadatos genérico para todo el repositorio. El prototipo contempla además la configuración del protocolo OAI-PMH de acuerdo a lo estipulado en las Directrices del Sistema Nacional de Repositorios Digitales (SNRD). La idea del prototipo fue permitir que instituciones con escasos recursos tanto humanos como tecnológicos pudieran disponer de un modelo ya diseñado para crear su propio repositorio institucional, en base a desarrollos previos existentes en otras instituciones. Para ello se redactó un instructivo de instalación, la descripción completa del producto, el diccionario de datos que describe los metadatos utilizados y un tutorial de uso y digitalización básica. El prototipo se publicó en el sitio web del CNG en acceso abierto, a fin de que los posibles interesados pudieran descargarlo.

En relación a las encuestas realizadas, se recibieron un promedio de 50 respuestas en cada una (2009, 2011 y 2016), donde el nivel de capacitación de los usuarios fue aumentando en cada oportunidad, en concordancia con las exigencias sobre las prestaciones disponibles en el software, la especificidad en las dificultades encontradas y los requerimientos de soluciones a la comunidad de desarrolladores.

Entre las razones que determinaron la elección del software por sobre otros disponibles, predominó la facilidad de instalación y uso, su bajo nivel de requerimiento tecnológico, la falta de recursos humanos disponibles en cantidad y calidad para abocarse a la tarea de creación y mantenimiento de las bibliotecas digitales y los repositorios

institucionales, y la evidente complejidad de instalación y configuración de otros sistemas existentes.

Finalmente, entre las dificultades y debilidades con Greenstone detectadas en las encuestas, se pueden mencionar: la pequeña comunidad de desarrolladores, acotada a la Universidad de Waikato y consultores asociados, la escasa cantidad de documentación disponible en la Wiki del software (<http://wiki.greenstone.org>) y una limitada oferta de consultores especializados en el país.

4 ESTADO ACTUAL EN ARGENTINA: CARACTERIZACIÓN DE LAS IMPLEMENTACIONES

Para conocer el uso actual de Greenstone en Argentina, se realizó un relevamiento específico de los sitios desarrollados con esa tecnología. Como base se tomó la información disponible en el CNG, que posee una lista cronológica de las bibliotecas digitales existentes. A esas, se sumaron otras que no figuraban y a la vez se eliminó un sitio que ya no estaba vigente. Cabe aclarar que el listado incluye bibliotecas digitales hechas con Greenstone que son de acceso público y están disponibles a través de la web. No se incluyeron los desarrollos de uso interno que poseen algunas instituciones³.

En total se localizaron 25 implementaciones activas y en producción en instituciones de diverso tipo en Argentina, las cuales pueden verse en una tabla en el Anexo.

Si observamos el tipo de aplicación desarrollada, de la lista total se observa que en 16 casos se trata de repositorios institucionales digitales, 8 son bibliotecas digitales dedicadas a una temática en particular (ej. agrimensura) o a un tipo documental (ej. fotografías), y 1 caso es un repositorio digital de fuentes para la investigación.



3 Por ejemplo, el Archivo Digital de la Secretaría de Derechos Humanos del Ministerio de Justicia de la Nación, que cuenta con una biblioteca digital de uso interno, la cual posee alrededor de 3 millones de objetos digitales

Gráfico 1. Tipo de aplicação desenvolvida com Greenstone

Fuente: Relevamiento realizado ad-hoc para este trabajo

Considerando que Greenstone no es un software específico para el desarrollo de repositorios digitales institucionales, es llamativa la cantidad de instituciones que optaron por utilizarlo con esta finalidad, y a la vez se observa una interesante variedad. Aunque predominan las instituciones universitarias -tanto públicas como privadas-, también aparecen en menor medida organismos de ciencia y tecnología (CNEA e INTA), dependencias gubernamentales (Centro de Información y Documentación del Ministerio de Economía y Biblioteca Nacional Mariano Moreno), asociaciones civiles (Epimeleia y Bolsa de cereales), una asociación profesional (Colegio Profesional de Agrimensura) y una red de bibliotecas de institutos de investigación del área de ciencias sociales (CLACSO).

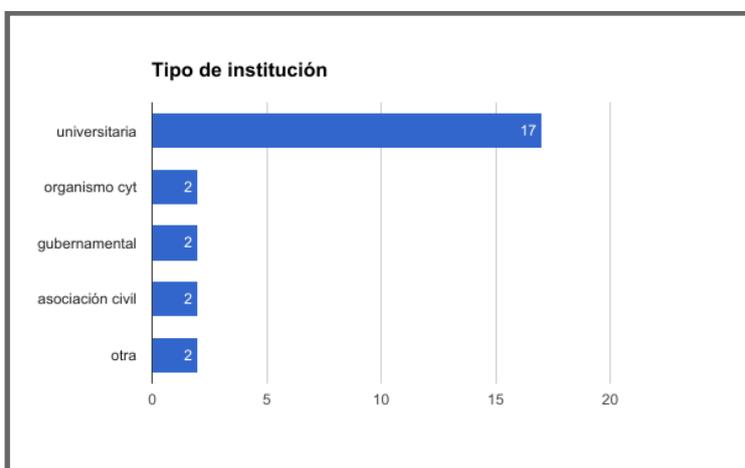


Gráfico 2. Tipo de institución que utiliza Greenstone

Fuente: Relevamiento realizado ad-hoc para este trabajo

Para dimensionar el tamaño de los desarrollos existentes, se registró la cantidad tanto de colecciones como de objetos digitales incluidos en ellas.

Respecto de las colecciones, que cada sitio organiza según distintos criterios y objetivos, se observó una nutrida diversidad, ya que si bien incluyen mayormente texto en formatos diversos, existen también importantes colecciones de fotografías, medallas, manuscritos y documentos históricos, recortes periodísticos, audios y materiales audiovisuales. En total se registraron 112 colecciones en todos los sitios.

Respecto de las colecciones textuales, predomina la inclusión de producción científico tecnológica (tesis, artículos, libros y partes de libros, informes técnicos, revistas, ponencias y actas de congreso, entre otras) en el caso de las instituciones que usan Greenstone como repositorio. Pero en muchos casos se observó la inclusión de documentos textuales de tipo institucional, como son los programas de materias, cursos y seminarios que se dictan, los planes de estudio, proyectos y memorias institucionales, documentos de archivo, actas de reuniones y otros.

Para estimar la cantidad de objetos digitales incluidos, se realizaron búsquedas con el comando “site” y “filetype” en el motor de búsqueda de google⁴. Se encontraron alrededor de 480.000 archivos en total, de los cuales 81.200 son archivos con formato pdf.

Los dos cuadros que siguen muestran el tamaño de los desarrollos existentes según la cantidad de objetos digitales contenidos. La tabla de archivos *pdf* refiere principalmente a documentos textuales, ya que es el formato más utilizado, aunque se sabe que también se utilizan otros (doc, docx, ppt, pptx, txt, etc.) que en esta cuenta no se incluyeron. En cambio, la tabla de archivos totales muestra la totalidad de documentos, aunque las cantidades finales podrían estar un poco sobrevaloradas ya que incluye toda la tipología de archivos:

Tabla 1. Cantidad de archivos registrados por desarrollo (pdf y totales)

Archivos pdf	Cantidad	%		Archivos totales	Cantidad	%
Entre 0 y 99	5	20%		Entre 0 y 99	1	4%
Entre 100 y 999	9	36%		Entre 100 y 999	4	16%
Entre 1000 y 9999	9	36%		Entre 1000 y 9999	8	32%
Entre 10000 y 99999	2	8%		Entre 10000 y 99999	12	48%
Total	25			Total	25	

Fuente: Relevamiento realizado ad-hoc para este trabajo

Luego, para dimensionar estas cantidades, se relevó la cantidad total de objetos digitales reunidos en el Portal del Sistema Nacional de Repositorios Digitales (SNRD) dependiente del Ministerio de Ciencia, Tecnología e Innovación Productiva (Mincyt), que sumaba en ese momento un total de 97.360 registros⁵ recolectados mediante el protocolo OAI-PMH de las alrededor de 40 instituciones adheridas. Cabe señalar que para integrar dicho portal, además de contar con un repositorio institucional en producción y haber adherido al sistema, es necesario cumplir con las Directrices del SNRD y exponer los metadatos de los objetos digitales en el formato allí requerido.

La tabla que sigue muestra el aporte de registros de las instituciones adheridas al SNRD según el tipo de software utilizado por cada una de ellas:

⁴ Salvo en el caso de Prohuerta y Adala, debido a que sus colecciones parecen no haber sido indizadas por el motor de búsqueda. En esos casos se contó manualmente la cantidad de objetos disponibles en el sitio.

⁵ Abril 2017

Tabla 2. Cantidad de repositorios y documentos aportados al SNRD según el software utilizado

Software utilizado	Cantidad de repositorios	Cantidad de documentos	% de documentos aportados
DSpace	12	64626	66%
Greenstone	5	22802	23%
OJS	1	6068	6%
Eprints	1	2153	2%
Otros ⁶	3	1711	2%
Total	22	97360	

Fuente: Relevamiento realizado ad-hoc para este trabajo

Como puede observarse allí, a pesar de que la mayoría de los repositorios adheridos (12) utiliza el conocido software DSpace, aportando el 66% de los registros, los cinco repositorios desarrollados con Greenstone aportan una cantidad significativa: el 23% del total, en el segundo lugar de la tabla.

5 GREENSTONE VERSIÓN 3: CARACTERÍSTICAS, IMPLEMENTACIONES Y DIFERENCIAS CON LA VERSIÓN 2

Greenstone3 es un rediseño completo de la antigua versión de Greenstone2 que estaba programada en C++. Sin embargo Greenstone3 conserva muchos aspectos de su manejo, configuración y servicios. La interfaz del bibliotecario (GLI⁷), que es la herramienta que permite crear, parametrizar y construir colecciones, sigue funcionando igual e inclusive los comandos para realizar tareas por consola siguen siendo similares en ambas versiones. Los cambios profundos son sobre la tecnología interna y los nuevos servicios, y no sobre las herramientas de manejo y configuración. Este nuevo desarrollo conserva todas las ventajas de Greenstone2: multiplataforma, altamente configurable y multilingüe (SPAN0, 2013).

Escrito en Java, el nuevo software está estructurado como una red de módulos independientes que se comunican mediante XML. Luego los documentos XML son

⁶En Otros se incluyen 2 repositorios que tienen desarrollos propios y 1 hecho con el software de gestión bibliotecaria PMB (<http://www.sigb.net/>)

⁷Greenstone Librarian Interface

convertidos a HTML mediante tecnología XSL que se presentan con el contenedor de servlets TOMCAT. Las colecciones intercambiables por el protocolo SOAP hacen que la información contenida en ellas no sólo sea legible por humanos, sino también por máquinas, mediante la utilización de web services. En el cuadro que sigue se explican con mayor detalle las tecnologías usadas:

Tabla 3. Tecnologías aplicadas en GS3

Tecnología	Características
JAVA	Oracle, Java Research License (JRL) Uno de los lenguajes de programación más populares en uso de aplicaciones cliente-servidor web. Solo requiere una máquina virtual Java (JVM) para su ejecución.
XML	Desarrollado y Recomendado por W3C. Lenguaje de marcado de documentos legible por máquina. Muy utilizado para el intercambio información estructurada entre plataformas.
XSL	Estándar de la W3C Para la transformación de documentos XML utilizando reglas. La unión de XML y XSLT permite separar contenido y presentación.
SOAP	Bajo el auspicio de la W3C Protocolo para el intercambio de información estructurada por medio de intercambio de datos en XML.
SOLR	Apache Software Foundation. Código abierto bajo licencia Apache 2.0 Motor de búsqueda desarrollado en Java. Incluye búsqueda en texto completo, resaltado de términos, facetado de los resultados, indexación en tiempo real, escalabilidad y tolerancia a fallos.
TOMCAT	Apache Software Foundation. Código abierto bajo licencia Apache 2.0 Contenedor de servlets.

Fuente: Elaboración propia

Las mejoras y nuevos servicios incorporados en Greenstone3 pueden resumirse en los siguientes puntos:

- a) URL limpias: Greenstone3 utiliza urls más simples que consisten en términos separados por barras inclinadas en orden jerárquico y ya no se incorporan los engorrosos parámetros de la versión 2

- b) Mayor personalización de la interfaz: El programa genera los datos de la página en XML que se convierten en HTML mediante XSLT. Esto ofrece completa libertad al momento de definir la visualización para cada página de la biblioteca digital
- c) Búsqueda cruzada entre colecciones: La búsqueda entre las distintas colecciones ahora es automática e incluye la posibilidad de ordenar los resultados por diferentes criterios como en la búsqueda de cada colección
- d) Facetado de los resultados de búsqueda: Con la implementación del motor de búsquedas Solr, los resultados de una búsqueda se pueden facetar (filtrar) con distintos criterios parametrizables a nivel de colección
- e) En una misma instalación varias bibliotecas digitales con diferentes interfaces: Una única instalación de Greenstone3 puede conformar varias Bibliotecas Digitales o conjuntos de colecciones y cada una de ellas se puede mostrar con un número ilimitado de interfaces (Ej. interfaz para celular y tablets, para impresión etc)
- f) RSS: El servicio de RSS está activado automáticamente en cada colección
- g) Recolección de documentos: Cada colección tiene la posibilidad de activar la herramienta “Cesta”, que consiste en poder seleccionar documentos de las colecciones para la exportación de sus registros
- h) Fragmentos de búsqueda: En los resultados de la búsqueda se puede mostrar los fragmentos de texto que contiene el término de búsqueda que fue encontrado en el documento

6 SITIOS DESARROLLADOS CON LA VERSIÓN 3

Aunque la versión 3 comenzó a programarse en 2002, fue recién en 2016 que desde el sitio oficial del software se puso a ésta como la versión principal de descarga, anunciando asimismo que si bien la versión 2 se mantendría vigente, ya no se continuaría con su desarrollo.

A pesar de lo reciente del cambio, en el ámbito local se han iniciado algunos desarrollos incipientes y prometedores a la vez, dos de los cuales se comentan a continuación.

El primero de ellos es ARCAS, el repositorio de fuentes de interés para la investigación inaugurado en 2014, que combina el uso de la versión 3 de Greenstone junto con un sistema de gestión de contenidos (CMS) (Plone) para manejo de los usuarios y la actualización de los contenidos del sitio mediante un formulario web. En este caso, y al tratarse de un desarrollo relativamente nuevo, se decidió desde el inicio hacerlo con la versión 3, por lo que no hubo un trabajo de migración o adecuación de colecciones preexistentes.



Figura 2. Página de inicio de ARCAS

Fuente: Recuperado de <http://arcas.fahce.unlp.edu.ar> el 28-4-2017.

Cada colección está a cargo de un investigador de la Facultad que administra los contenidos, las exhibiciones y el grupo de trabajo académico. Una de las premisas del desarrollo fue reducir al mínimo la dependencia del personal de biblioteca en el mantenimiento del portal y de las fuentes. Para dar solución a este punto se implementó un formulario de edición de metadatos, accesible internamente desde el gestor de contenidos Plone o desde la visualización de los documentos en Greenstone previo chequeo de autenticación. De este modo los investigadores editan la descripción de los documentos existentes o agregan nuevos a la colección.

El formulario en Plone utiliza la tecnología de web service de Greenstone para obtener los datos de Colección, Serie, Subserie y Obra, en el caso de la edición de los metadatos de un documento existente. Una vez elegida la obra se carga la descripción actual y el investigador modifica el registro. La nueva versión será revisada por el personal técnico de la biblioteca y actualizada en el portal público.



Figura 3. Formulario de edición de metadatos en ARCAS
Fuente: Recuperado de <http://arcas.fahce.unlp.edu.ar> el 28-4-2017.

La colección de los manuscritos de Manuel Puig, en particular, reúne imágenes digitalizadas que han sido organizadas y descritas por el equipo de investigación a cargo. En este caso, se utilizó la implementación de la visualización asincrónica de las imágenes que componen cada manuscrito, desarrollo incluido en Greenstone3 para la representación de imágenes paginadas.

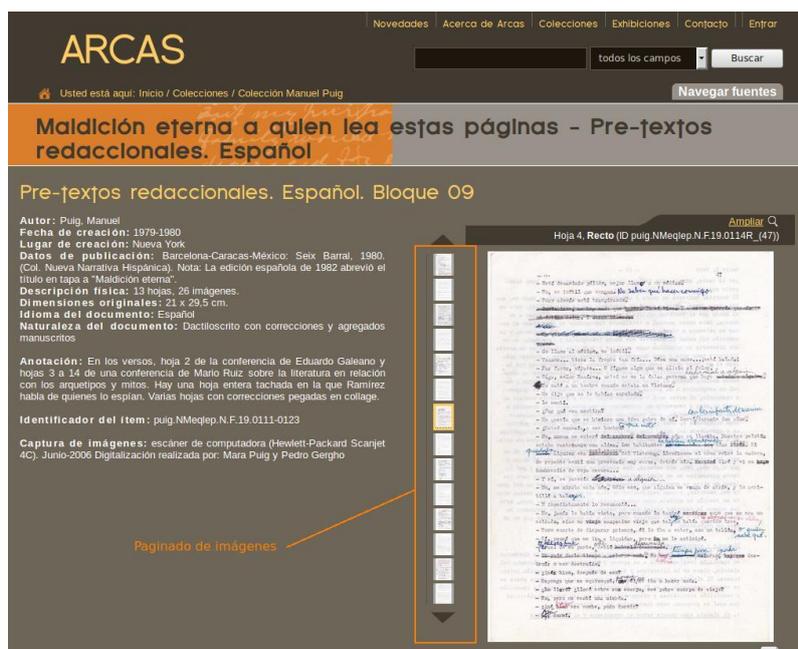


Figura 4. Paginado de imágenes en ARCAS
Fuente: Recuperado de <http://arcas.fahce.unlp.edu.ar> el 28-4-2017.

El segundo sitio comentado en este trabajo es la nueva versión⁸ de la Biblioteca Digital de la Facultad de Ciencias Exactas y Naturales de la UBA, que partió del sitio preexistente y realizó la migración de todas sus colecciones a la versión 3.

Esto implicó generar un nuevo archivo de configuración para cada colección, ya que el archivo que parametriza las colecciones en Greenstone 2 no es aplicable a Greenstone 3, que incluye tecnología xslt.

Por otro lado se generó una nueva visualización de la interfaz incluyendo los estilos de Bootstrap que produce sitios renderizables.

The screenshot shows the 'Biblioteca Digital' interface. At the top, there's a blue header with the site name and a 'Colecciones' dropdown. Below it, a breadcrumb trail reads 'Biblioteca Digital / Articulos / Documento'. The main content area displays a document record for 'Marchi, M.C.; Bilmes, S.A.; Ribeiro, C.T.M.; Ochoa, E.A.; Kleinke, M.; Alvarez, F. "A comprehensive study of the influence of the stoichiometry on the physical properties of TiOx films prepared by ion beam deposition" (2010). Journal of Applied Physics. 108(6)'. It includes download and access options, social media icons, and an abstract section. On the right, a search sidebar is visible with 'Buscar en:' and 'Listar por:' sections.

Figura 5. Vista del registro de un artículo en la Biblioteca Digital de FCEN-UBA

Fuente: Recuperado de <http://digital-beta.bl.fcen.uba.ar/> el 28-4-2017.

Esta implementación incorpora los nuevos servicios de Greenstone 3: Búsqueda cruzada entre todas las colecciones, el facetado de los resultados de la búsqueda y muestra fragmentos del texto con el término de búsqueda destacado.

⁸La versión beta de este desarrollo puede consultarse en: <http://digital-beta.bl.fcen.uba.ar/> [fecha de consulta 28-4-2017].

The screenshot displays the search results interface of the FCEN-UBA Digital Library. At the top, the header includes the library name and a 'Colecciones' dropdown. Below, the search criteria are shown: 'Buscar en: all fields' and 'cell*'. The results list shows 821 documents. A list of related terms is provided, such as 'cell (18103)', 'cella (3)', 'celladhesive (1)', etc. The first result is by Bermúdez Moretti, M., et al. (1995) about GABA uptake in a yeast strain. The second result is by Bildner, A.G., et al. (2013) about galectins at the fetomaternal interface. The third result is by Perone, M.J., et al. (2009) about suppression of autoimmune diabetes. Annotations with green boxes and arrows point to: 1) 'Indicador de frecuencia de términos' pointing to the list of terms; 2) 'Filtrar los resultados' pointing to the search filter; 3) 'Segmento de texto con término de búsqueda destacado' pointing to the highlighted word 'Cells' in the abstract of the third result; 4) 'Filtrar resultados por:' pointing to the filters on the right, including 'Descarga', 'Año', and 'Tipo de Documento'.

Figura 6. Vista de la pantalla de resultados de búsqueda en la Biblioteca Digital de FCEN-UBA
Fuente: Recuperado de <http://digital-beta.bl.fcen.uba.ar/> el 28-4-2017.

7 REFLEXIONES FINALES Y PERSPECTIVAS DE FUTURO

Para finalizar, en este apartado exponemos algunas reflexiones que nos surgen a partir de los desarrollos observados y la situación actual.

En primer lugar, nos interesa destacar que los propósitos y características originales del software -que se mantienen en la nueva versión- permitieron que instituciones con escasos recursos tanto humanos como tecnológicos pudieran desarrollar bibliotecas digitales de todo tipo a partir de un producto sencillo, fácil de instalar y de usar, altamente configurable y adaptable a la mayor variedad de requerimientos posibles. Este fue el motivo inicial por el que el software se difundió tan ampliamente en nuestra región y en nuestro país, y la razón por la cual existe hoy una importante cantidad de bibliotecas y repositorios digitales en plena producción que lo utilizan.

En el ámbito de los repositorios institucionales, tanto a nivel global como en nuestro país hubo inicialmente un período de prueba y experimentación que se caracterizó por la implementación de repositorios en diversos softwares. En una segunda etapa, la

experimentación de software fue desplazada por la adopción de un software que se volvió de uso mayoritario. Actualmente las instituciones rara vez pasan por una fase de experimentación y en cambio eligen esa plataforma ya consolidada y más difundida, en un contexto dado por la mayor capacidad de difusión que tiene el software de uso mayoritario y las instituciones que lo respaldan, aunque el producto no cubra todas sus necesidades o se adecue a sus características. Si bien hoy DSpace se presenta como una solución que dice contemplar todas las funcionalidades posibles requeridas para este tipo de desarrollos y al tratarse de un software de código abierto, siempre es posible introducir adaptaciones y mejoras, claramente se trata de un producto que requiere de conocimientos avanzados en programación para su eventual modificación. Esto restringe la posibilidad de que instituciones que no cuenten con personal informático dedicado puedan utilizarlo, o bien que acepten la versión estándar, resignando la posibilidad de adaptar el uso del software o su diseño final a las necesidades específicas y tercerizando las actividades técnicas.

Un aspecto de no poca importancia es la ductilidad de Greenstone para representar en la interfaz del usuario muy diversas formas organizativas de las instituciones y de su producción intelectual. Greenstone no impone una metáfora para la representación de la información, sino que ésta es libremente diseñada por los responsables de los repositorios o las bibliotecas digitales. Por el contrario, encontramos en los repositorios creados con DSpace, una misma y reiterada organización de la información que se nuclea bajo el término “Comunidades”. A veces este término refiere a colecciones, otras a dependencias de una institución y otras posibilidades. El resultado es el ocultamiento de formas organizativas que no pueden ser cabalmente representadas en su variedad. Las instituciones fuerzan así una forma de mostrar su producción que no es fiel a los elementos ni a las formas en que el ciclo de producción intelectual se realiza.

En este sentido, sostenemos la importancia de mantener cierto nivel de diversidad en la utilización y desarrollo de software que permita atender a los requerimientos y situaciones específicas que se presentan en los distintos ámbitos, sin pretender homogeneizar o imponer una tecnología como la única elección posible, con excepción del cumplimiento de los protocolos y estándares necesarios. Con frecuencia esta pretensión de uniformidad y estandarización lleva a la imposición de ciertas tecnologías que en definitiva lo que promueven es la predominancia de un determinado producto por sobre el resto, eliminando a los supuestos competidores, y con ello, también la posibilidad de que otros grupos e iniciativas crezcan y se desarrollen, generando a su vez un círculo productivo de creación de conocimiento.

Por otra parte, al observar en detalle los desarrollos existentes realizados con Greenstone, vemos una importante riqueza y variedad, tanto en el tipo de colecciones y objetos incluidos, como en su diseño. Puede observarse asimismo la reutilización y aprovechamiento de recursos y bases de datos previas. Todo eso fue posible gracias a que la facilidad del software permitió que muchos bibliotecarios/as se animaran a desarrollar y

gestionar sus propias colecciones. En las instituciones en las que se utilizó Greenstone, el personal afectado realizó un fuerte aprendizaje y adquirió nuevas habilidades en el uso de tecnologías de la información que le permitieron liderar sus propios desarrollos.

Una de los mayores desafíos que enfrenta Greenstone actualmente, es el desarrollo de una comunidad con mayor participación de desarrolladores. Actualmente el perfil predominante de la comunidad es de integrantes de bibliotecas, archivos e instituciones de la memoria autodidactas en el uso de herramientas informáticas que les permiten satisfacer las necesidades de las instituciones a las que pertenecen y sus usuarios. La facilidad de implementación del software y los escasos recursos necesarios para su uso crearon una comunidad de diversos perfiles profesionales con diferentes niveles de conocimientos técnicos que fueron el fuerte en el comienzo de su historia. El desarrollo de nuevas funciones y la corrección del código quedó casi en su totalidad a cargo del grupo de desarrollo original, con pocas intervenciones de desarrolladores del resto del mundo. Muchas de las necesidades fueron asistidas e implementadas en cada nueva versión, pero no es suficiente frente al creciente avance de nuevas tecnologías.

Desde el Centro Nacional de Promoción de Greenstone en Argentina hacemos un llamado a la reflexión sobre estas cuestiones e invitamos a las instituciones de la región y del país que lideran las redes de repositorios digitales a acompañar y sostener los desarrollos existentes, apoyando y fortaleciendo a la comunidad de usuarios de Greenstone y manteniendo una base de interoperabilidad mínima que permita la diversidad e integración de acuerdo a las posibilidades de cada uno.

Y finalmente también invitamos a sumarse para conocer y explorar las potencialidades que ofrece la versión 3 de Greenstone para el desarrollo nuevas aplicaciones y la migración de los sitios existentes.

APÊNDICE A LISTA DE IMPLEMENTACIONES DE GREENSTONE EN ARGENTINA

TI=tipo de institución / TS=tipo de sitio / COL=cantidad de colecciones

#	Nombre del sitio	Institución	TI	URL	TS	COL	pdf	total files
1	Repositorio institucional	Instituto Universitario de Ciencias de la Salud Fundación H. A. Barceló	UN	http://www.barcelo.edu.ar/greenstone/cgi-bin/library.cgi	RI	4	458	3360

2	Repositorio institucional	Instituto Universitario y Hospital Italiano de Buenos Aires	UN	http://trovare.hospitalitaliano.org.ar/	RI	5	406	2990
3	Fauba digital	UBA. Facultad de Agronomía	UN	http://ri.agro.uba.ar/	RI	2	2030	19300
4	Repositorio	UBA. Facultad de Ciencias Económicas	UN	http://bibliotecadigital.econ.uba.ar/	RI	7	3560	17400
5	Biblioteca Digital	UBA. Facultad de Ciencias Exactas y Naturales	UN	http://digital.bl.fcen.uba.ar/	RI	4	7620	18700
6	Memoria institucional	UBA. Facultad de Medicina	UN	http://www.bibliomedicinadigital.fmed.uba.ar/medicina/cgi-bin/library.cgi	RI	2	383	94500
7	Biblioteca Digital	Universidad Católica Argentina	UN	http://bibliotecadigital.uca.edu.ar/	RI	6	6050	38300
8	Archivo fotográfico	Universidad Católica de Córdoba	UN	http://bibdigital.uccor.edu.ar/gsd/cgi-bin/library.cgi	BD	1		15500
9	Repositorio Digital Institucional	Universidad de Buenos Aires	UN	http://repositorioubas.sisbi.uba.ar/	RI	5	1320	25300
10	RDD Repositorio Documental y de Datos	Universidad Nacional de Avellaneda	UN	http://rdd.undav.edu.ar/	RI	5	30	458
11	Repositorio digital de acceso abierto	Universidad Nacional de La Pampa	UN	http://www.biblioteca.unlpam.edu.ar/greenstone/cgi-bin/inicio.htm/	RI	5	4750	26800
12	Repositorio Digital Institucional José María Rosa	Universidad Nacional de Lanús.	UN	http://www.repositoriojmr.unla.edu.ar/	RI	1	307	2230

13	Archivo periodístico del exilio argentino en México y Recursos digitales externos	Universidad Nacional de Lanús.	UN	http://www.unla.edu.ar/greenstone/cgi-bin/library.cgi?site=localhost&a=p&p=about&c=archived&l=es&w=utf-8	BD	2	9140	27000
14	Repositorio institucional	Universidad Nacional de San Martín	UN	http://ri.unsam.edu.ar/	RI	1	151	7760
15	Biblioteca Digital Arq. Hilario Zalba	UNLP. Facultad de Arquitectura y Urbanismo	UN	http://bdzalba.fau.unlp.edu.ar/greenstone/cgi-bin/library.cgi	RI	6	179	1330
16	Memoria académica	UNLP. Facultad de Humanidades	UN	http://www.memoria.fahce.unlp.edu.ar/	RI	10	19800	89400
17	Arcas	UNLP. Facultad de Humanidades	UN	http://arcas.fahce.unlp.edu.ar/	RF	3	2	269
18	Red de bibliotecas virtuales de ciencias sociales	CLACSO	otra	http://www.biblioteca.clacso.edu.ar/	RI	3	19300	39100
19	Biblioteca digital de agrimensura	Consejo Profesional de Agrimensura de la Provincia de Buenos Aires	otra	http://www.bibliotecacpa.org.ar/greenstone/cgi-bin/library.cgi	BD	6	1280	8730
20	Biblioteca digital histórica	CNEA. Centro de información Eduardo Savino	OCyT	http://bdcies.cnea.gov.ar/	BD	3	551	1260
21	Biblioteca Digital Prohuerta	Instituto Nacional de Tecnología Industrial	OCyT	http://prohuerta.inta.gov.ar/biblioteca/	BD	6		294
22	Acervo digital anotado de literatura argentina	Biblioteca Nacional Mariano Moreno	GOB	http://adala.fahce.unlp.edu.ar/	BD	1		42

23	Biblioteca digital	Ministerio de Economía. Centro de documentación informativa	GOB	http://cdi.mecon.gob.ar/greenstone/cgi-bin/library.cgi	BD	14	287	404
24	Epimeleia Biblioteca para la persona	Asociación civil Epimeleia	AC	http://bibliotecaparaalapersona-epimeleia.com/greenstone/cgi-bin/library.cgi	BD	6	431	30300
25	Biblioteca Digital	Bolsa de cereales	AC	http://bolsadecereales.com.ar/greenstone/cgi-bin/library.cgi	RI	4	3090	8050
	Total de colecciones y objetos digitales					112	81125	478777

REFERÊNCIAS

CENTRO NACIONAL DE PROMOCIÓN DE GREENSTONE EN ARGENTINA [sitio web]. Recuperado de: <http://cng.fahce.unlp.edu.ar/> [fecha de consulta 27-4-2017]

CYRANEK, G., (ed.). **Greenstone** : Un software libre de código abierto para la construcción de bibliotecas digitales. Experiencias en América Latina y el Caribe. Montevideo [UY] : UNESCO, 2010.

GREENSTONE DIGITAL LIBRARY SOFTWARE [sitio oficial]. Recuperado de: <http://www.greenstone.org/> [fecha de consulta 29-4-2017]

GREENSTONE WIKI. Recuperado de: <http://wiki.greenstone.org/doku.php> [fecha de consulta 29-4-2017]

SPANNO, D. **#Greenstone**: migrando hacia una biblioteca digital 3.0. Charla organizada por la comunidad de práctica SocialBiblio y transmitida el 9 de abril 2013. Recuperado de: <http://www.socialbiblio.com/materiales/greenstone-migrando-biblioteca-digital-30>

PICHININI, Mariana. **Prototipo para repositorios institucionales utilizando Greenstone** – versión beta. 2012. Colaboradoras: Marcela Coria y Carolina Unzurrunzaga. Recuperado desde: <https://pictobonaerense.wordpress.com/publicaciones/> [29-4-2017]

UNESCO. **Curso MERCOSUR sobre la Construcción de Bibliotecas Digitales**. Realizado en la Biblioteca Nacional de Uruguay. Montevideo, 29 de agosto al 2 de septiembre 2005.

