

PONENCIA

Proyectos conjuntos entre grupos de investigación y bibliotecas para la preservación digital y la difusión. El caso de un archivo de fuentes primarias de autores destacados

Claudia González^{1,2}; Mónica G. Pené^{1,2}; Carolina Unzurrunzaga¹

1 Facultad de Humanidades y Ciencias de la Educación de la Universidad Nacional de La Plata (FaHCE/UNLP)

2 Instituto de Investigaciones en Humanidades y Ciencias Sociales (IdIHCS/UNLP-CONICET)

Mail de correspondencia: cgonzalez@fahce.unlp.edu.ar

Resumen: El interés por preservar y difundir digitalmente los materiales de autores destacados de la cultura de un país es un buen motivo de reunión de investigadores humanistas y Bibliotecas universitarias. A partir del ejemplo de gestación del proyecto Archivo ARCAS de fuentes primarias de autores destacados, se muestra la riqueza de interacciones posibles en un escenario marcado por la sinergia que produce el desarrollo de proyectos conjuntos. Se exponen los antecedentes y se presentan los actores intervinientes. Se detallan las principales decisiones que la FaHCE (UNLP) / IdIHCS (UNLP-CONICET) y su Biblioteca han debido tomar teniendo en cuenta su organización, la disponibilidad de recursos humanos y la base tecnológica existente para que el proyecto pueda prosperar. Luego, a un nivel más técnico, se consignan las principales características del prototipo generado y de su primera fase de implementación. Finalmente, se concluye enunciando los desafíos a futuro que el proyecto demandará, considerando la construcción de una plataforma integrada de fuentes primarias que contemple las necesidades de grupos de investigación con perfiles heterogéneos.

Palabras clave: Grupos de investigación; Bibliotecas universitarias; Repositorios de fuentes primarias; Acceso abierto; Archivos de escritores.

Abstract: The interest in preserving in digital format and dissemination materials produced by prominent authors of a country provides good reasons for gathering researchers in the humanities in university libraries. Using the example of the development of the ARCAS archive project of primary sources of notable authors, it is shown the fruitfulness of possible interactions in a scenario characterized by the synergy produced by the development of cooperative projects. The

antecedents of the project and the intervening actors are presented. The main decisions that FaHCE (UNLP) and IdHCS (UNLP-CONICET) and their library had taken are described in a detailed way, taking into account their organization, the availability of human resources and the existing technological base that is necessary for the success of the project. At a more technical level, the main characteristics of the generated prototype and the first phase of implementation are presented. Finally, as a conclusion, the future challenges demanded by the project are stated, which require the construction of an integrated platform of primary sources that considers the needs of research groups with heterogeneous profiles.

Keywords: Research Groups; University Libraries; Repositories of Primary Sources; Open Access; Writers's Archive.

1. Introducción

La reunión de documentos en el desarrollo de proyectos de investigación es algo habitual en el ámbito de las Humanidades y de las Ciencias Sociales, ya que es muy difícil pensar la construcción de nueva ciencia sin que ésta se apoye en el conocimiento previamente publicado y difundido que existe en nuestra sociedad. Pero, ¿qué sucede cuando estos documentos son inéditos y su acceso es restringido? ¿Cómo proceden los investigadores en esta etapa de la investigación cuando los documentos no se encuentran en una institución (archivo, biblioteca o museo) que les facilite el acceso y la recuperación de la información contenida en ellos?

Generalmente, cuando un grupo de investigación detecta fuentes primarias¹ que aportan información relevante para sus estudios, inicia un arduo trabajo que comienza con la gestión de permisos para actuar sobre la colección. Este trabajo continúa con un relevamiento exhaustivo que implica diversas tareas de organización y preparación de los materiales, que incluye, en la mayoría de los casos, la generación de imágenes digitales y una mínima descripción.

Una vez que la investigación ha finalizado -o se encuentra en su etapa final- puede vislumbrarse cierto interés por parte de los miembros del grupo en que el trabajo realizado con los documentos trascienda los límites de los resultados arrojados por el propio proyecto. Esta situación suele darse cuando los investigadores toman conciencia de que las colecciones con

¹ Fuentes que contienen información original, de primera mano, que no han recibido ningún tipo de tratamiento. Definición tomada de López Yepes, José (ed.). *Diccionario enciclopédico de ciencias de la documentación*. Madrid : Síntesis, 2004.

las que han estado trabajando son de interés cultural para una región o país, incluso en muchos casos constituyen parte de su patrimonio documental, que se encuentra oculto en colecciones privadas, a veces desprotegido o en riesgo, e innumerables veces inaccesible (Pené, 2012). En ese momento es cuando se plantean la necesidad de su preservación -sobre todo en digital- y su difusión a través de la web² y se les presentan inquietudes, necesidades y limitaciones para cumplir el cometido.

La construcción de archivos digitales en abierto, que cumplan con los estándares requeridos y con una apariencia acorde con los actuales productos de información web, implica poseer una cantidad de conocimientos y de habilidades, además de involucrar una movilización de recursos considerable. Es en estos aspectos donde las capacidades del propio grupo de investigación, por lo general, resultan insuficientes y es oportuno entonces plantear un escenario de trabajo interdisciplinario con profesionales del área de las bibliotecas y los archivos, así como también de la informática y del diseño.

En el caso particular de la biblioteca universitaria, a lo largo de su historia siempre ha servido de apoyo al ámbito de la investigación. Recorriendo sus antecedentes puede distinguirse que en las décadas de 1950 y 1960, ante el crecimiento exponencial de la información científica, las bibliotecas fueron las que facilitaron las tareas de los investigadores al potenciar los servicios de resúmenes y de disseminación selectiva de la información. Luego, entre los años 1970 a 1990, el principal apoyo al investigador giraba en torno a la búsqueda de información en bases de datos especializadas y a la obtención de los textos completos correspondientes (Pérez Alvarez-Osorio, 1988). La explosión de la World Wide Web, a mediados de los '90, abrió nuevos horizontes para la difusión de información y comenzaron a gestarse en las bibliotecas proyectos de bibliotecas digitales de colecciones especiales y repositorios institucionales, los que, actualmente, se han transformado en herramientas importantes para la preservación y la difusión de la producción de investigación de las universidades (Hernández Pérez, Rodríguez Mateos y Bueno De la Fuente, 2007). La experiencia adquirida por este tipo de biblioteca en la gestación de estos proyectos y los equipos de especialistas que han quedado conformados es lo que permite aseverar que los actores y los conocimientos necesarios para generar archivos digitales se encuentran también en la biblioteca, presentándose como un escenario de oportunidades para una fructífera

² Esto dependerá tanto de la naturaleza y el estado de las colecciones como de los acuerdos y permisos que el grupo haya conseguido con el propietario, familia o albacea de las fuentes analizadas.

interacción entre grupos de investigación y bibliotecas.

En este marco, se presenta la experiencia suscitada en la Facultad de Humanidades y Ciencias de la Educación de la Universidad de La Plata (FaHCE/UNLP³) entre distintos grupos de investigación literarios del Instituto de Investigaciones en Humanidades y Ciencias Sociales (IdIHCS/UNLP-CONICET), que funciona dentro de dicha Facultad, el grupo de investigación de Bibliotecología que trabaja temas relacionados con el acceso abierto desde 2011⁴ y la Biblioteca Profesor Guillermo Obiols, que atiende a la comunidad investigadora. Esta experiencia es un ejemplo de trabajo interdisciplinario para la creación de un archivo digital de fuentes primarias, ARCAS, a fin de asegurar la preservación digital y la difusión en acceso abierto de los documentos que son fuentes de investigaciones.

2. Antecedentes de un contexto favorable

Desde hace décadas, investigadores literarios pertenecientes a la FaHCE/UNLP vienen desarrollando diferentes tipos de investigaciones vinculadas a lo que se ha dado en denominar “crítica genética”, definida por Élica Lois (1995: 402) como “el estudio de la prehistoria de los textos literarios, es decir, el desciframiento, análisis e interpretación de los papeles privados de un autor”. Por regla general, el objeto de análisis de la crítica genética suelen ser los documentos escritos, de preferencia los manuscritos, que agrupados en conjuntos coherentes conforman “la huella visible de un proceso creativo”. Durante su trabajo, los investigadores han relevado y descubierto diversos papeles de escritores, alojados en casas particulares, debiéndose ocupar en primera instancia de su organización, digitalización y descripción para luego proceder a su estudio. Tal es el caso concreto del grupo dirigido por la Dra. Graciela Goldchluk, que en el primer año del proyecto “Archivos de la memoria escritural del Cono Sur: documentación, integración cultural, procesos textuales y teoría de la edición” (H395)⁵ se acercó a la Biblioteca de la Facultad para solicitar asesoramiento en cuanto a normas y pautas para la

³ En adelante, FaHCE-UNLP.

⁴ Banzato, Guillermo, director. (2009) El acceso abierto al conocimiento científico en Argentina / H540 (Proyecto de investigación) UNLP. FaHCE. Datos del proyecto disponibles en: <http://www.memoria.fahce.unlp.edu.ar/library?a=d&c=proyecto&d=Jpy527>

⁵ Goldchluk, Graciela Beatriz, director. (2004) Archivos de la memoria escritural del Cono Sur: documentación, integración cultural, procesos textuales y teoría de la edición / H395 (Proyecto de investigación) UNLP. FaHCE. Centro de Estudios de Teoría y Crítica Literaria - Departamento de Letras. Datos del proyecto disponibles en: <http://www.memoria.fahce.unlp.edu.ar/library?a=d&c=proyecto&d=Jpy324>.

descripción de documentos manuscritos correspondientes a la colección privada del escritor argentino Manuel Puig (1932-1990). Éste fue el primer hito que marcó el inicio de la relación/interrelación entre los investigadores literarios y la Biblioteca.

Varios años después, el mismo grupo de investigación contactó nuevamente a la Biblioteca. Conscientes del valor cultural que presentan los documentos con los que trabajan, entendiendo que era necesaria su preservación y difusión para las generaciones futuras, pero superados por las circunstancias y los conocimientos ausentes, los investigadores literarios se acercaron para solicitar la generación de un espacio institucional donde se pudieran alojar las versiones digitalizadas de los papeles relevados en el transcurso del proceso investigativo. Se manifestó entonces un cambio de actitud en los investigadores, que comenzaron a asumir así un rol activo como agentes que propugnan la preservación a largo plazo y la difusión vía web de aquellas fuentes primarias empleadas para sus propias producciones académico-científicas, pensando, por un lado, en que esos materiales constituyen parte de nuestro patrimonio documental y, por otro lado, que debe ser ampliamente difundido para que otros puedan estudiarlos. A finales de la década de 2000, motivados por dar visibilidad a los manuscritos de Puig, la Biblioteca de la FaHCE en coordinación con el grupo antes mencionado generó un espacio virtual dedicado al Archivo Digital Puig⁶, dentro del sitio web de la Biblioteca. Este primer espacio, a modo de borrador, sirvió además para ir aprendiendo el camino a seguir en este tipo de iniciativas, que requieren no sólo de tecnologías y técnicas específicas, sino también de la resolución de cuestiones legales vinculadas a los permisos para la difusión de fuentes en acceso abierto.

Con el transcurso del tiempo, el número de fuentes disponibles y de actores interesados en aportar documentos digitales se acrecentó, situación que dio inicio a un proceso de análisis para el diseño de una herramienta más adecuada para la gestión de este tipo de materiales. Revisando la bibliografía existente (Arano et al, 2011; Hernández-Pérez, 2013) y otras experiencias previas, surgió la idea de generar un repositorio o archivo digital de fuentes primarias, es decir, una base de datos en línea donde las fuentes fueran depositadas y descritas conforme a estándares que permitan su posterior identificación y recuperación por otros investigadores. Considerando la postura institucional de la FaHCE respecto del acceso abierto al conocimiento científico-académico, se pensó desde un comienzo en la generación de un repositorio abierto. Así, miembros del grupo de investigación bibliotecológico reunidos en el

⁶ <http://www.fahce.unlp.edu.ar/biblioteca/labiblioteca/archivo-digital-manuel-puig>.

proyecto "El movimiento de Acceso Abierto al conocimiento científico en la Argentina. Políticas y prácticas en torno a la investigación, las revistas académicas y los repositorios" (H-642) comenzaron a interactuar con personal de la Biblioteca en busca de información y herramientas para la concreción de los objetivos perseguidos. Las primeras reuniones estuvieron orientadas a delimitar el contenido del repositorio, para luego dar inicio a las actividades vinculadas a la elección del software de gestión y a los estándares y protocolos que debían considerarse para cumplir con las pautas internacionales en materia de preservación, descripción e interoperabilidad de los objetos digitales. Las siguientes reuniones tuvieron como objetivo delimitar la política de funcionamiento del repositorio, involucrando a las autoridades de la Facultad y estableciendo normativa institucional que reglamentara las actividades propuestas.

3. Diseñando un proyecto perdurable

Realizar proyectos que por sus características se enmarcan dentro de las Humanidades Digitales implica, tal como lo sostiene Daniel V. Pitti (2008), poseer familiaridad con el tema de investigación y con las tecnologías disponibles. Sin embargo, que el grupo de investigación conozca sobre el tema, y que la Biblioteca de su Universidad disponga y domine las tecnologías y los recursos necesarios para realizar la edición y la publicación web de productos informativos, no resulta suficiente. Tal como se detalla a continuación, en el caso ARCAS fue necesario trabajar en la articulación de diversos aspectos de tipo normativo y organizacional a fin de concebir un proyecto que se pueda sostener a largo plazo.

3.1. Institucionalización del archivo

Desde el momento en que un grupo de investigación decide trabajar en la organización de determinadas fuentes, dedica tiempo de sus integrantes y dinero de sus propios subsidios para emprender las primeras tareas. Si la interacción entre el grupo de investigación y la Biblioteca se reduce a la solicitud de asesoramiento en temas de digitalización, conservación o la descripción de recursos a nivel general, no parece ser necesaria más que la buena disposición de los profesionales bibliotecarios. Sin embargo, si la propuesta es llevar adelante un Archivo Digital disponible en la web, la envergadura del proyecto excede por lo general los recursos a disposición del propio grupo y compromete los de la Biblioteca que se encuentra abocada a otras tareas. Por lo tanto, en el caso ARCAS, como puede serlo en cualquier

proyecto similar, lograr el apoyo al interior de la institución universitaria fue uno de los pasos fundamentales para obtener los recursos materiales y humanos necesarios.

El proyecto se formalizó mediante una resolución que fue aprobada por el Consejo Directivo de la FaHCE/UNLP (Res. 1308/2012⁷), en la que se crea un “Archivo digital de fuentes de interés para la investigación, en el cual se registrarán, preservarán y difundirán en acceso abierto, los manuscritos y demás documentos reunidos y procesados por los grupos de investigación que manifiesten interés en hacerlo”, a la vez que se indica que el Archivo se iniciará con documentos de autores (ARCAS). En su texto se hace mención además a dos aspectos importantes. Por un lado, establece que la creación de cada fondo documental deberá gestarse en el ámbito de los Centros e Institutos de Investigación que funcionan dentro del ámbito de la propia Facultad, quién además se encargará de designar un responsable académico por cada uno de ellos. Por otro lado, encomienda a la Biblioteca la tarea de desarrollo y mantenimiento del Archivo y la hace responsable de las tareas de coordinación entre las partes.

De esta manera, se hacen explícitos los roles de las partes, disponiendo una frontera entre los aspectos académicos a resolver por los Centros y las cuestiones de tipo operativo que atenderá la Biblioteca, garantizando a su vez la existencia de un cierto nivel de uniformidad en el tratamiento de las colecciones, impuesto por la propia naturaleza de un proyecto que aspira a ser una plataforma para múltiples colecciones. Por otra parte, y lo más importante, es que la formalización de la existencia del Archivo y el contar con responsables asignados, permite tener interlocutores al momento de salir a obtener nuevos fondos dentro de la propia institución, así como contar con una figura que reúne a un grupo de interesados frente a posibles financiamientos externos.

3.2. Conformación de los grupos de trabajo

Los recursos humanos involucrados en este tipo de proyectos poseen características bien diferentes. Por un lado, tal como ya se ha mencionado, el grupo de investigación conforma su propio equipo en el momento en que decide trabajar las fuentes. Si bien el número de personas puede variar, lo integran por lo general algún investigador de carrera, becarios doctorales y alumnos avanzados realizando trabajos de grado. Bajo el liderazgo del investigador

⁷ Resolución Nro.1308/12. Creación del Archivo digital de fuentes de interés para la investigación. (2012). Recuperado a partir de: <http://www.memoria.fahce.unlp.edu.ar/normativa/no.50/50.pdf>.

principal, este equipo debe resolver la obtención del permiso para trabajar con los papeles, la organización física -si es que no estaba ya resuelta al momento de iniciar las actividades-, la digitalización in-situ y la organización de los archivos digitales con un registro de metadatos mínimo.

La Biblioteca, por su parte, también conforma su equipo de trabajo en el momento que decide coordinar este proyecto. Debe reunir a profesionales y técnicos de diferentes disciplinas para cubrir aspectos muy distintos. En primer término es necesario contar con conocimientos en propiedad intelectual, licencias libres y negociación de derechos, ya que será necesario evaluar las condiciones legales de una colección antes de ser incorporada al archivo. En segundo término, es necesario contar con conocimientos de archivística y documentación digital para diseñar la estructura del archivo, seleccionar los estándares de descripción apropiados y generar el perfil de aplicación. Estos profesionales serán los encargados también de delinear el plan de preservación digital. En tercer término, es necesario contar con conocimientos de edición web, y éste involucra dos aspectos interrelacionados. Por un lado se deben manejar con solvencia herramientas tecnológicas para gestionar/publicar productos web, a la vez que estos conocimientos deben complementarse con conceptos de usabilidad, diseño orientado al usuario y diseño en comunicación visual. Por último, es necesario contar con recursos humanos que estén entrenados en la catalogación de recursos, con habilidades de interpretación, normalización, manejo de software y ajuste a ritmos de producción propios del trabajo en serie.

En el caso de ARCAS el grupo de trabajo está integrado por una archivera con conocimientos en derecho de autor, dos bibliotecarias especialistas en documentación digital, una bibliotecaria con manejo de software web, un diseñador en comunicación visual con conocimientos en usabilidad web, un técnico especializado en informática documental y una bibliotecaria para catalogación y control de la carga de metadatos. Todos ellos trabajando a tiempo parcial y en el caso del diseñador y del informático, corresponden a un servicio tercerizado. Pero además de los conocimientos técnicos mencionados, tanto la archivera como las bibliotecarias, son miembros de grupos de investigación en temáticas afines al proyecto, por lo que se suma a sus conocimientos profesionales, un entendimiento general de la dinámica de trabajo de los grupos de investigación y el compartir un lenguaje común.

3.3. Articulación de los procedimientos y flujos de información

La gestión de las autorizaciones, ya sea para trabajar sobre los papeles como para la publicación web en acceso abierto, son gestionadas por el investigador principal. Él es quien tiene acceso al autor, a su familia o albaceas. La Biblioteca comienza a intervenir cuando el investigador le trae la autorización correspondiente o bien, previamente, si le solicitara asesoramiento sobre las modalidades de autorización posibles. Se considera que no será necesaria la participación de otras autoridades, ya que más allá de los aspectos legales, el investigador ha sabido establecer un vínculo de confianza con la contraparte que percibe de él un tratamiento académico y responsable para con toda la obra. En este escenario, la intervención de terceras partes puede no ser conveniente.

El trabajo de digitalización y descripción de los objetos digitales implica la interacción entre los grupos a nivel de asesoramiento mutuo y coordinación de las tareas de control de los productos que se van obteniendo en las diferentes etapas. Si el grupo de investigación no sabe cómo abordar el proceso de organización de los papeles, la digitalización y/o la descripción mínima de los recursos, podrá pedir asesoramiento para la compra de equipamiento, para la sistematización del flujo de trabajo, para la elaboración de guías de procedimientos ad-hoc o la utilización de herramientas informáticas de registro. Por otro lado, las partes deberán pautar la forma y condición en la que se entregarán los materiales para el Archivo. La Biblioteca se encargará del control de calidad de los digitales y controlará y completará, si fuera necesario, el registro descriptivo. Al llegar a esta etapa es necesario que existan canales de comunicación fluidos de manera que se produzca un *feed-back* que garantice calidad en los productos.

En el caso ARCAS, el diseño del prototipo implicó un trabajo intenso hacia al interior de la Biblioteca. En un primer momento se abordó el trabajo de estructura, definiendo el perfil de la aplicación que requirió el compromiso de las bibliotecarias documentalistas y la archivera. Luego se procedió a realizar la bajada del esquema a la aplicación informática seleccionada (ver el siguiente apartado) y realizar la descripción de algunos ejemplos. Recién con estos elementos a disposición, fue posible hacer una propuesta inicial a los investigadores responsables de las colecciones, de manera que teniendo a la vista un prototipo de sistema factible, les fuera posible pensar las características que desearían que tuviera la interfaz web del Archivo. La interacción con el equipo de diseño web se dió en torno a la elaboración y comentarios sobre propuestas de maquetado, las cuales nunca perdieron de vista los diferentes intereses a los que el producto final debía responder. Esto es: un producto destinado al mundo académico, pero que a su vez busca interesar al público general; y que además debe estar

desarrollado con las tecnologías *open source* que usa la Biblioteca.

4. Características del prototipo

Como se mencionó anteriormente, ARCAS es un repositorio de fuentes primarias en donde se depositan papeles digitalizados de artistas. En este apartado se expondrán las características más sobresalientes del prototipo, el cual se encuentra actualmente en etapa de prueba. La exposición se centra en aspectos técnicos relacionados con el software elegido, el perfil de aplicación definido para el registro de metadatos dentro del prototipo y su implementación, la organización documental en el repositorio y la interoperabilidad de éste con otros sistemas de información. Se incluye también una descripción de su interfaz web.

4.1. Software

El producto diseñado se sustenta en Greenstone (http://www.greenstone.org/index_es), versión 2.85. Este software, de código abierto y distribución gratuita, permite generar y distribuir colecciones digitales que pueden conformar cualquier tipo de repositorio digital. Ha sido desarrollado por la Universidad de Waikato (Nueva Zelanda) y distribuido por la UNESCO y la ONG Human Info (Bélgica). Destacan, entre sus particularidades, las siguientes:

- Servidor OAI integrado que cumple con el protocolo OAI-PMH. Los metadatos de los objetos publicados pueden ser cosechados por cualquier servicio de harvesting. El programa permite mapear cualquier esquema de metadatos con Dublin Core (DC), que es el utilizado por este protocolo (Pichinini, 2010).
- Interfaz del bibliotecario (GLI), un workflow que permite la gestión de las colecciones. Facilita la configuración, el registro de información, el almacenamiento de archivos, el diseño de colecciones y la definición de índices para el buscador y de clasificadores para la navegación de las colecciones.
- Interfaz del lector, desde donde el usuario accede a las colecciones, a través de un navegador web.
- Gestor de esquemas de datos, interfaz “*metadata set editor*”, que permite generar o modificar esquemas de metadatos.

El software se eligió por ser una herramienta en la que se tiene experiencia previa: el repositorio institucional⁸ lo utiliza desde 2006 y los participantes del proyecto de investigación ya

⁸ Memoria Académica, repositorio institucional de la FaHCE/UNLP. Disponible en:

han trabajado con éste en otras ocasiones, y principalmente por la gestión que admite de los esquemas de metadatos: permite generar perfiles de aplicación propios y trabaja con el concepto de herencia en la asignación de metadatos.

4.2. Esquema de metadatos

Para afrontar la descripción de los objetos digitales se ha diseñado un perfil de aplicación híbrido, que integra elementos procedentes de diversos esquemas de metadatos existentes. Los esquemas que se seleccionaron, por cumplir con los requerimientos detectados y con los criterios de preservación de aplicación internacional, son:

- *Metadata Encoding and Transmission Standard* (METS) (<http://www.loc.gov/standards/mets/>), un estándar de metadatos de tipo descriptivos, administrativos y estructurales. Este esquema sirve para describir objetos digitales y también para conjugar distintos esquemas a fin de armar perfiles de aplicación híbridos. En el esquema definido, éste es el que lo estructura.
- *Encoded Archival Description* (EAD) (<http://www.loc.gov/ead/>), un estándar propio de la disciplina archivística lo suficientemente preciso para describir los tipos documentales a incluir. Este esquema posee una estructura lógica y jerárquica de un instrumento de descripción de archivo lo que permite difundir, acceder y navegar la información descriptiva disponible (Peis & Ruiz-Rodríguez, 2004). La información descriptiva está organizada en el perfil de aplicación según lo indica la norma. Contempla los niveles de descripción archivística e incluye en cada uno de éstos los elementos que permiten su descripción física y de contenido.
- *Preservation Metadata: Implementation Strategies* (PREMIS) (<http://www.loc.gov/standards/premis/>), un estándar de metadatos de preservación muy utilizado a nivel mundial. Los metadatos PREMIS introducen en el esquema la descripción técnica de las versiones de los documentos y describen los procesos a los que éstos han sido sometidos y los actores que han intervenido.
- *Metadata for Images in XML Standard* (MIX) (<http://www.loc.gov/standards/mix/>) , un estándar específico para la descripción técnica de imágenes digitales. En el perfil, estos metadatos optimizan la descripción a realizar de las imágenes depositadas.

La implementación de este perfil en el software ha requerido un trabajo especial para lograr la aplicación del concepto de herencia, a partir del cual cada documento hereda los datos descriptivos de los niveles superiores para completar la información de su registro. La incorporación de este concepto permite asegurar la consistencia en la asignación de metadatos y también disminuir la cantidad de datos a ingresar y a almacenar ya que no es necesaria la repetición de las propiedades definidas en un nivel más alto en todas las agrupaciones documentales de los niveles subordinados (UNE-ISO 23081-2:2009).

4.3. Aspectos de preservación

Se han seleccionado formatos estandarizados para la generación de las imágenes de preservación (TIFF), para las imágenes de difusión (JPEG) y para los documentos descargables (PDF). En esta etapa se han delineado pautas de preservación que luego deberán hacerse explícitas en un Plan documentado. Se ha previsto un doble respaldo de los objetos digitales y de la totalidad del sistema que actualmente se encuentran alojados en el servidor de la Biblioteca. Por un lado se duplicará la información en el sistema de almacenamiento de datos de la Facultad que se encuentra en el mismo edificio. Por otro lado, se encuentra bajo consideración el lugar de almacenamiento remoto, que fuera de la ciudad garantizará la preservación ante desastres naturales. Para la comprobación de integridad se aplicará un algoritmo de *checksum* que se utilizará para almacenar el dato en el momento de la ingesta, y se ha programado un script que permite ser corrido con regularidad para realizar las comparaciones de los valores actuales con los valores almacenados para comprobar que el contenido no ha sido alterado.

4.4. Organización documental

Las colecciones del repositorio se organizan por criterio de pertenencia. Está previsto que exista una colección particular para los papeles de cada uno de los autores. Siguiendo los criterios archivísticos, las colecciones se agrupan por niveles que reflejan la organización dada a los materiales por el propio autor y/o por el grupo de investigación. Esta estructura queda registrada, a su vez, en los metadatos específicos definidos en el perfil de aplicación. Por el momento, en esta etapa de prueba, se encuentra disponible parte de la colección digitalizada del escritor argentino Manuel Puig y del artista plástico platense Edgardo A. Vigo. Se están gestionando los permisos necesarios para incluir documentos de otros autores/artistas que son

trabajados por los investigadores de la Facultad.

4.5. Interoperabilidad

El prototipo desarrollado es interoperable con cualquier sistema de información a través del protocolo OAI-PMH, que exige el uso del esquema de metadatos Dublin Core (DC) sin calificar como medio para alcanzar la interoperabilidad. Para ello se ha mapeado el perfil de metadatos creado con el esquema DC. Actualmente se está trabajando para compartir los documentos almacenados en ARCAS con la *Plataforma Interactiva de Investigación para las Ciencias Sociales* (PLIICS)⁹.

4.6. Interfaz web

Para la puesta en línea del prototipo, además de Greenstone, se ha utilizado Plone, un gestor de contenidos basado en el paradigma de objetos. La página principal de ARCAS y las correspondientes a la presentación y el acceso a cada colección se sustentan en este sistema, lo que permite su gestión distribuida y compartida entre la Biblioteca y los grupos de investigación y la incorporación de información dinámica. Entre las particularidades del entorno web del prototipo pueden mencionarse las siguientes:

- Contiene dos tipos de buscadores: uno general en la página principal, que permite realizar búsquedas transversales en todas las colecciones, y otros específicos en la página de cada colección, que posibilita la realización de búsquedas en una colección en particular.
- Cada colección permite una navegación por los nombres de las obras, por los niveles de la colección y, también, cronológicamente.
- Los documentos se muestran en pantalla junto con su descripción y su licencia de uso. Se ofrece también una versión descargable en formato pdf.
- Brinda acceso a la producción relacionada con los documentos fuentes archivados. Se enlazan tanto estudios depositados en el repositorio institucional como cualquier otro que el coordinador de cada colección considere pertinente.

⁹ Más información de PLIICS disponible en: <http://web.conicet.gov.ar/web/conicet.trabajar.investigacion/-ciencias-sociales>.

5. Desafíos futuros

La experiencia expuesta aquí destaca la importancia del trabajo interdisciplinar para la correcta preservación digital y difusión web de las fuentes primarias que se han utilizado en las investigaciones. También pone de relieve el rol que pueden jugar las bibliotecas universitarias en la construcción de plataformas de publicación de materiales digitales que cumplan con los estándares internacionalmente aceptados. Resalta, asimismo, el hecho de que la organización bibliotecaria haya gestado con anterioridad los repositorios institucionales, ya que llega a esta instancia con una importante experiencia previa y además, en términos funcionales, presenta la posibilidad de enlazar las fuentes tratadas con los estudios académicos que se han realizado. Esto permite generar recursos para la investigación con un mayor sentido de integración.

Si bien se ha iniciado el camino, se plantea como desafío el hacer crecer el Archivo aplicando una política proactiva hacia la inclusión de nuevas colecciones sin descuidar su tratamiento y la jerarquía de los autores incluidos a fin de convertirlo en un lugar de referencia para la cultura regional. Queda aún avanzar en el desarrollo de una intranet que facilite las tareas *on-line* de todos los actores involucrados, ya que es imprescindible minimizar esfuerzos y liberar tiempo para las verdaderas tareas de producción investigadora.

Si bien el Archivo de autores ARCAS es el primer proyecto que iniciamos en relación con fuentes y datos primarios de investigación, entendemos que la experiencia adquirida y el entusiasmo puesto en el tema nos llevará a abordar otras tipologías documentales y otras disciplinas dentro de las Humanidades y Ciencias Sociales.

6. Referencias bibliográficas

- Arano, S., Martínez, G., Losada, M., Villegas, M., Casaldàliga, A., & Bel, N. (2011). La comunidad «Recursos y datos primarios» de la Universitat Pompeu Fabra: los repositorios institucionales como infraestructuras científicas: estudio de caso. *Revista española de Documentación Científica*, 34(3): 385-407. doi: 10.3989/redc.2011.3.834.
- Hernández-Pérez, T. (2013). Datos abiertos y repositorios de datos: ¿nuevo reto para los bibliotecarios? *Notas ThinkEPI 2013*, 4 de febrero de 2013. Recuperado el 25 de mayo de

2013 a partir de <http://www.thinkepi.net/datos-abiertos-repositorios-datos-nuevo-reto-bibliotecarios>.

- Hernández Pérez, T., Rodríguez Mateos, D., & Bueno De la Fuente, G. (2007). Open Access: el papel de las bibliotecas en los repositorios institucionales de acceso abierto. *Anales de Documentación*, 10:185-204. Recuperado el 25 de mayo de 2013 a partir de <http://hdl.handle.net/10201/4017>.

- Lois, E. (1995). Amado Alonso, precursor de la crítica genética. *CAUCE, Revista de Filología y su Didáctica*, (18-19): 401-408. Recuperado el 25 de mayo de 2013 a partir de http://cvc.cervantes.es/literatura/cauce/pdf/cauce18-19/cauce18-19_26.pdf.

- Peis, E., & Ruiz-Rodríguez, A. A. (2004). EAD (Encoded Archival Description): Desarrollo, estructura, uso y aplicaciones. *Hipertext.net*, (2). Recuperado el 25 de mayo de 2013 a partir de <http://www.upf.edu/hipertextnet/numero-2/ead.html>.

- Pené, M. G. (2012). De la caja a la vidriera : Los manuscritos de escritor en acceso abierto. Ponencia presentada en 1er. Congreso de la Delegación Argentina de la Asociación de Lingüística y Filología de América Latina, 21 al 23 de marzo de 2012, La Plata. Recuperado el 25 de mayo de 2013 a partir de http://www.memoria.fahce.unlp.edu.ar/trab_eventos/ev.2467/ev.2467.pdf.

- Pérez Alvarez-Ossorio, J. (1988). *Introducción a la información y documentación científica*. Madrid: Alhambra.

- Pichinini, M. (2010). Argentina: Los repositorios institucionales como vía de acceso abierto al conocimiento científico. En G. Cyranek (ed.). *Greenstone: Un software libre de código abierto para la construcción de bibliotecas digitales. Experiencias en América Latina y el Caribe*. Montevideo: UNESCO. Recuperado el 25 de mayo de 2013 a partir de <http://www.unesco.org.uy/ci/fileadmin/comunicacion-informacion/Greenstone-paraweb.pdf>.

- Pitti, Daniel V. (2004). Designing sustainable projects and publications. En S. U. Schreibman (eds.). *Digital humanities*. Oxford: Blackwell.

- UNE-ISO 23081-2:2009. Información y documentación. Procesos de gestión de documentos. Parte 2: Elementos de implementación y conceptuales. Madrid: AENOR.

