

Esta conferencia fue presentada en las V Jornadas sobre Imagen, Cultura y Tecnología, en la Universidad Carlos III de Madrid, España, el 3 de julio de 2006. Los términos de su uso se establecen por medio de una licencia libre de la agrupación Creative Commons. Quedan permitidas su reproducción y distribución para fines no comerciales, siempre que se cite debidamente la fuente. Además, sólo se permite su distribución bajo los términos de la misma licencia. Los detalles de ésta pueden consultarse en el sitio <<http://creativecommons.org/licenses/by-nc-nd/2.0/deed.es>>. Para citar en trabajos académicos esta conferencia o las ideas que desarrolla, aplican las mismas reglas conocidas de ese ámbito.

METADATOS TRANSFORMADOS.
ARCHIVOS DIGITALES, LA WEB SEMÁNTICA Y
EL NUEVO PARADIGMA DE LA CATALOGACIÓN

ANDREW RUSSELL GREEN
Instituto de Investigaciones Dr. José María Luis Mora

Resumen

El proyecto “Patrimonio cultural e histórico latinoamericano” genera metodologías para la construcción de archivos, coloca fondos digitales en línea, estudia fenómenos sociales con base en ellos, y desarrolla un sistema informático que se ejecuta en un servidor Web y surte a los usuarios las páginas con dichos fondos. El proyecto se concibe como una investigación interdisciplinaria en archivología, ciencias sociales e informática. Esta ponencia describe el proyecto y presenta algunas reflexiones teóricas que ha suscitado, referidas sobre todo al marco interdisciplinario del mismo y el uso de las representaciones del conocimiento en la catalogación. Se afirma que con el uso cada vez más frecuente de esta tecnología surge un nuevo paradigma en las prácticas catalográficas.

Abstract

The project “Cultural and Historical Heritage of Latin America” aims to develop methodologies for archive creation, deploy on-line digital collections, study social phenomena by way of said collections, and develop a computer program that will run on Web servers and will provide users with the pages containing the digital collections. The project is conceived as interdisciplinary research in Archival Science, Social Science and Computer Science. This paper describes the project and presents some of the theoretical reflections it has generated, primarily regarding the project's interdisciplinary framework and the use of knowledge representations in cataloging. The author claims that the increasingly frequent use of that technology is causing a new paradigm in cataloging practice to form.

En el año 2002 el Laboratorio Audiovisual de Investigación Social¹ propuso crear una herramienta informática para difundir archivos fotográficos por Internet. En ese momento no anticipamos la cantidad y diversidad de problemas teóricos y filosóficos que enfrentaríamos en el transcurso de esa labor. Posteriormente se reformuló y se amplió el proyecto, que ahora lleva el nombre “Patrimonio cultural e histórico latinoamericano” (PCHLA), y se desarrollaron propuestas teóricas para sustentarlo. A continuación expongo algunos aspectos de dichas propuestas y las reflexiones que han suscitado, relacionadas sobre todo con el carácter interdisciplinario del proyecto y el uso de las representaciones del conocimiento en la catalogación.

Un punto de partida del proyecto fue la preocupación por recuperar el valor epistémico de las imágenes y fomentar su integración como fuentes primarias de investigación. Por medio de un estudio previo varios miembros del proyecto actual constataron que un obstáculo recurrente para la recuperación

¹ El Laboratorio Audiovisual de Investigación Social (LAIS) forma parte del Instituto de Investigaciones Dr. José María Luis Mora, institución pública mexicana adscrita al Concejo Nacional de Ciencia y Tecnología y ubicada en la Ciudad de México. Agradezco la asesoría del equipo del LAIS en la elaboración de esta ponencia.

de las imágenes como fuentes de información es la situación de los archivos fotográficos. No relataré todos los problemas que se detectaron en los archivos; me limito a señalar que una de las dificultades que a menudo enfrentan es la carencia de recursos, metodología y políticas para lograr una catalogación que facilite el uso de los materiales desde la investigación. Incluso en la Fototeca Nacional, archivo privilegiado por la cantidad de recursos que percibe en comparación con los demás del país, encontramos un catálogo con un número importante de errores y fichas semivacías. Por otra parte, al examinar las prácticas de uso de imágenes en las ciencias sociales se constató que los investigadores tienden a emplearlas más para adornar sus publicaciones que para analizar los procesos sociales.²

El proyecto PCHLA pretende incidir en esta problemática de distintas maneras: (1) desarrolla propuestas acerca de cómo construir archivos que faciliten la integración de los documentos fotográficos a los procesos de investigación; (2) genera un sistema informático para la creación de archivos digitales en línea con esta orientación y difunde dicho sistema bajo el modelo de *software* libre; (3) usa el sistema informático para colocar en línea fondos digitales que sirven de ejemplos de la metodología y enfoque archivístico que busca promover; y (4) estudia fenómenos sociales e históricos a partir de los fondos digitales mencionados y analiza los procesos de investigación que en ese marco se generen. Los cuatro productos—propuestas metodológicas, un sistema informático, fondos digitales en línea y conocimiento sociohistórico—se generan por medio de un proceso unificado de investigación interdisciplinario en archivología, informática y ciencias sociales. Finalmente, en un nivel filosófico, más allá de los objetivos de intervención en problemáticas determinadas, se parte de la tesis de que el significado profundo del patrimonio cultural e histórico se forja en la apropiación que hacen de él las diversas comunidades y sectores que conforman la sociedad, tesis que nos conduce a proponer el acceso libre a los bienes patrimoniales y al conocimiento social que se genera por medio de ellos.

En esta descripción resumida del proyecto PCHLA he enfatizado la preocupación por la relación entre archivos y conocimiento social porque las reflexiones que desarrollaré aquí se vinculan sobre todo con esta faceta del proyecto. Pero en sí el proyecto es más amplio; otros aspectos importantes del mismo sólo se mencionarán de paso—por ejemplo, la meta de fomentar la transparencia y la horizontalidad en las prácticas de investigación, y la valoración del conocimiento por su potencial como medio de transformación social.

Un elemento del proyecto que tiene mucho que ver con las reflexiones que siguen es el sistema informático que desarrollamos. Este *software* se ejecuta en un servidor Web; surte las páginas con los fondos digitales y proporciona funciones especializadas para el análisis de los mismos y para el intercambio de información entre investigadores. Para organizar los datos catalográficos el sistema

2 AGUAYO, Fernando y Lourdes Roca: “Estudio introductorio”, en Fernando Aguayo y Lourdes Roca (coords.), *Imágenes e investigación social*, México, Instituto Mora, 2005, pp. 9-28.

emplea la Web Semántica (WS), tecnología inspirada en la inteligencia artificial y la computación humanística que promete a mediano plazo aportaciones sustanciales: por medio de ella se espera lograr un manejo flexible y complejo de los metadatos, implementar funciones novedosas para la investigación social, y conjuntar en un mismo sitio Web una diversidad de fichas catalográficas. A la fecha hemos logrado una versión inicial del sistema, el cual hemos llamado “el Pescador” (haciendo analogía entre las redes para pescar y una red semántica que “pesca” los significados). Se han resuelto diversos problemas informáticos relacionados con el uso de la WS, como mencioné se ha avanzado en el desarrollo de las bases teóricas del proyecto, y se ha colocado en línea una primera colección digital. (Como anexo se incluye una descripción de las funciones actuales y futuras del Pescador, así como imágenes de pantalla de dicha colección.)

La interdisciplinariedad

En un sentido pragmático el carácter interdisciplinario del estudio se debe a la diversidad de las preguntas de investigación y actividades relacionadas con la construcción de los archivos digitales. Desde una perspectiva teórica el estudio es interdisciplinario en tanto rebasa las fronteras que se han erigido entre las disciplinas académicas, reorganiza las propuestas de algunas de ellas de una manera que permite su integración, y construye teorías y metodologías sin considerar dichas fronteras. En tanto, se busca construir un marco teórico-metodológico general, base que fundamentará todos los procesos del proyecto sin importar la disciplina a la que tradicionalmente pertenecen.

La elaboración de dicho marco general es el primer problema teórico no resuelto del proyecto. En este punto el único avance obtenido a la fecha es la siguiente propuesta: el eje rector del marco teórico-metodológico general debe ser la generación de conocimiento social e histórico y de mecanismos que permitan el acceso a él. Esto implica que la investigación social (concebida como práctica que encierra distintas posibilidades de transformación social, las cuales debe intentar realizar) amplíe sus fronteras y “se coma” a la archivología y la informática. En otras palabras, las prácticas y teorías archivísticas e informáticas se someten a la meta general de generar conocimiento sociohistórico y a los fundamentos teóricos y filosóficos de la ciencia social crítica. Pero en este proceso la investigación social, si bien mantiene intacto su núcleo duro, también queda trastocada: se adapta para poder “digerir” a las otras disciplinas.

Por ejemplo: la comunicación científica juega un papel central en los procesos investigación. Someter la investigación en informática al objetivo de crear conocimiento social implica entender a la computadora como un medio de organización, almacenamiento y transmisión de ideas que, al igual que el papel y la palabra hablada, modifica los mecanismos sociales de creación de conocimiento.³ Al diseñar

3 Una concepción parecida de la computadora ha sido desarrollada por el campo de la computación humanística. Véase

mecanismos de procesamiento digital debemos preguntarnos cómo éstos impactarán en las metodologías de investigación y las interacciones entre investigadores. Pero la informática, al mismo tiempo de integrarse a este marco general, impacta en las demás disciplinas involucradas en él; por ejemplo: tanto la noción de conocimiento social como la de datos catalográficos deben enfrentar los nuevos planteamientos acerca de la información, el significado y la comunicación desarrollados desde la informática.

Algo similar ocurre con la archivología—no obstante que se trata de una disciplina más cercana a la ciencia social que la informática. Esto se constata en la reformulación que se hizo del concepto de catalogación: en el marco interdisciplinario se entienden los materiales a catalogar como vestigios de procesos sociales. Por lo tanto, la información acerca de ellos debe construirse por medio de un estudio social. Desde esta perspectiva el llenar las fichas de un catálogo no es una labor técnica ni de capturista, sino de investigación, y como tal debe apoyarse en una diversidad de fuentes, fundamentarse en una reflexión crítica y explicitar tanto los supuestos que la enmarcan como la metodología que utiliza. Además es un trabajo que nunca termina, ya que los estudios acerca de la sociedad jamás alcanzan un punto final. Pero por otra parte la archivología, al igual que la informática, no sólo queda asimilada en el marco general, sino también impacta en las demás áreas del proyecto: por ejemplo resulta imprescindible desarrollar las nociones de relevancia y reutilización de la información con base en las propuestas de la archivología.⁴ Y esta disciplina nos recuerda que cuando generamos archivos sigue siendo una pregunta abierta el cómo concebir su utilización en estudios sociales con propósitos y bases mutuamente incompatibles.

Además de la necesidad de seguir elaborando de manera sistemática esta propuesta de marco teórico-metodológico general, el proyecto PCHLA enfrenta problemas prácticos causados por su carácter interdisciplinario. En concreto, la mayoría de los integrantes del proyecto se formaron en una disciplina determinada u otra. ¿Cómo lograr una comunicación profunda y detallada acerca de sus diversas actividades cuando la comprensión de éstas normalmente requiere de varios años de estudio? Una posible respuesta es: para nuevas formas de investigación, nuevos estudios. Pero resulta mucho más fácil decirlo que ponerlo en práctica.

Como ejemplo de la diversidad de actividades que el proyecto implica, presento a continuación una descripción (simplificada) del proceso que se llevó a cabo para la generar el primer fondo digital que colocó en línea el proyecto: el catálogo de Marcas de Fuego de la Benemérita Universidad Autónoma de Puebla. (Las marcas de fuego son un tipo de sello que aparece en los libros antiguos en México; probablemente fueron marcas de propiedad.)

McCARTY, Willard: “Humanities computing”, versión preliminar del artículo para *The Encyclopedia of Library and Information Science*, New York, Dekker, 2003. Fecha de consulta: 30 de mayo de 2006
<<http://www.cch.kcl.ac.uk/legacy/staff/wlm/essays/encyc/>>

4 Sobre este tema véase BUDD, John M.: “Relevance: Language, Semantics, Philosophy”, *Library Trends*, vol. 52, núm 3, (invierno 2004), Baltimore, John Hopkins University Press, pp. 447-462.

Generar conocimiento histórico es el objetivo central del catálogo, que podrá usarse para analizar cuáles libros fueron adquiridos y leídos por quiénes en Nueva España; esto podrá contribuir a la construcción de la historia de las ideas. Mencionamos esto a modo de ejemplo; sin embargo podemos imaginar estudios que aprovecharían este catálogo desde una gran diversidad de enfoques. Al diseñarse la ficha de las marcas de fuego se prevén los probables usos del catálogo y las informaciones que éstos necesitarán. Los criterios archivísticos, sigilográficos y heráldicos también influyen en este proceso. Luego el llenar la ficha requiere de diversos conocimientos de la historia y de los libros antiguos. Ahora, las fotografías que se toman de los libros son bienes digitales y deben crearse y gestionarse de acuerdo con parámetros archivísticos referentes a las imágenes de ese tipo. El manejo de los libros durante el proceso de digitalización se somete a las normas para el cuidado de libros antiguos. El traslado del catálogo al sistema computacional y la programación del mismo sistema se hace a partir de una investigación en informática, la cual, además de alimentarse de diversos trabajos recientes en esta disciplina, retoma ideas del ámbito de diseño industrial. Se proyectan las funciones del sistema con base en un análisis de las posibilidades que ofrece la tecnología para manipular los datos cualitativos y de las metodologías de investigación social por medio de las cuales podría aprovecharse esta fuente. Se dota al sistema de un diseño gráfico que, evidentemente, toma en cuenta el cuerpo de conocimientos que ha generado ese campo. Los textos explicativos que acompañan al catálogo describen el contexto social de la producción y uso de los bienes del catálogo y resumen el proceso de generación del mismo; para ello refieren conocimientos de casi todas las disciplinas mencionadas. Finalmente, en su concepción global este proceso integrara consideraciones teóricas, filosóficas y sociales acerca de la construcción de archivos digitales en línea.⁵

Reitero que estas reflexiones acerca de la interdisciplinariedad del estudio no son más que una propuesta parcial e inacabada. Todavía hace falta abordar numerosas facetas del problema. Por ejemplo: ¿Como conjuntar en el análisis nuestro entendimiento de los aspectos pragmáticos y teóricos de la interdisciplinariedad? Y ¿cómo integrar a la propuesta teórica las ideas de autores como Boaventura de Sousa y Emmanuel Wallerstein, quienes han estudiado la relación entre las ciencias sociales, la racionalidad y los procesos de desigualdad y exclusión social? (Esperemos que las respuestas que encontremos a preguntas como éstas nos ayuden a seguir construyendo una teoría y práctica interdisciplinaria—en lugar de obligarnos a desmantelar lo ya elaborado sobre esta cuestión.)

La Web Semántica y las representaciones del conocimiento

Otro aspecto del proyecto que ha sido objeto de una investigación teórica importante es el sistema informático y la tecnología que emplea, conocida como *Web Semántica* (WS). Esta tecnología permite estructurar los datos de una manera más flexible y más compleja que en los sistemas de información tradicionales. Esencialmente, la WS organiza los datos que se almacenan en una máquina de modo que un

5 LABORATORIO AUDIOVISUAL DE INVESTIGACIÓN SOCIAL: *Proyecto “Patrimonio cultural e histórico latinoamericano”*, texto inédito, 2006. Participaron en la elaboración del catálogo en línea investigadores de diversas instituciones; en la página <<http://www.marcasdefuego.buap.mx>> aparecen los créditos completos del mismo.

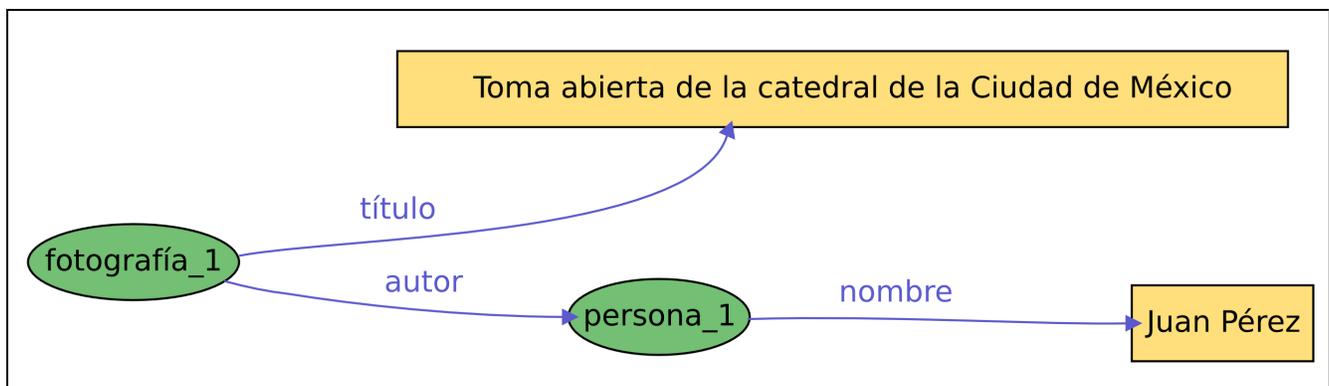
mayor porcentaje de los significados puedan ser manejados y comunicados automáticamente. Con este sistema, en vez de almacenar los datos en un conjunto de tablas (como sería el caso en una base de datos tradicional) la información queda registrada en una telaraña compleja de significados vinculados entre sí.

En la inteligencia artificial las estructuras de este tipo se llaman *representaciones de conocimiento* (RC); un conjunto de datos organizados con base en la WS constituye una RC. Nótese que no estoy hablando de la rama de inteligencia artificial que busca crear máquinas con voluntad propia, sino la que genera sistemas para el procesamiento de información compleja acerca de cualquier aspecto del mundo—por ejemplo, información sobre temas cotidianos que los seres humanos manejamos sin mayor esfuerzo, pero que generalmente queda fuera del alcance de los sistemas computacionales. Las RC almacenan y (aún de manera limitada) pueden procesar automáticamente información de este tipo, la cual las ciencias sociales llama “cualitativa” (para distinguirla de la información “cuantitativa”). En la informática las *ontologías* son un tipo de RC.

A continuación explicaré algunos aspectos del uso de las RC para gestionar los metadatos y describiré varios problemas y tendencias que se desprenden de ello.

Exploremos el tema mediante ejemplos. La Figura 1 es una representación gráfica de un pequeño fragmento de una RC que codifica el catálogo de una fototeca imaginaria.

Figura 1. Un fragmento de una RC hipotética



Esta imagen muestra el fragmento de la RC sin modificaciones. Ahora bien, evidentemente los datos no pueden presentarse ante los usuarios del catálogo en este formato, sino deben reorganizarse para adquirir un aspecto que facilite su comprensión. Las figuras 2 y 3 muestran posibles formatos de presentación de estos datos; despliegan más o menos el mismo contenido, pero organizado de otra manera.

Figura 2. Una reorganización de la información contenida en el fragmento de la RC de la Figura 1

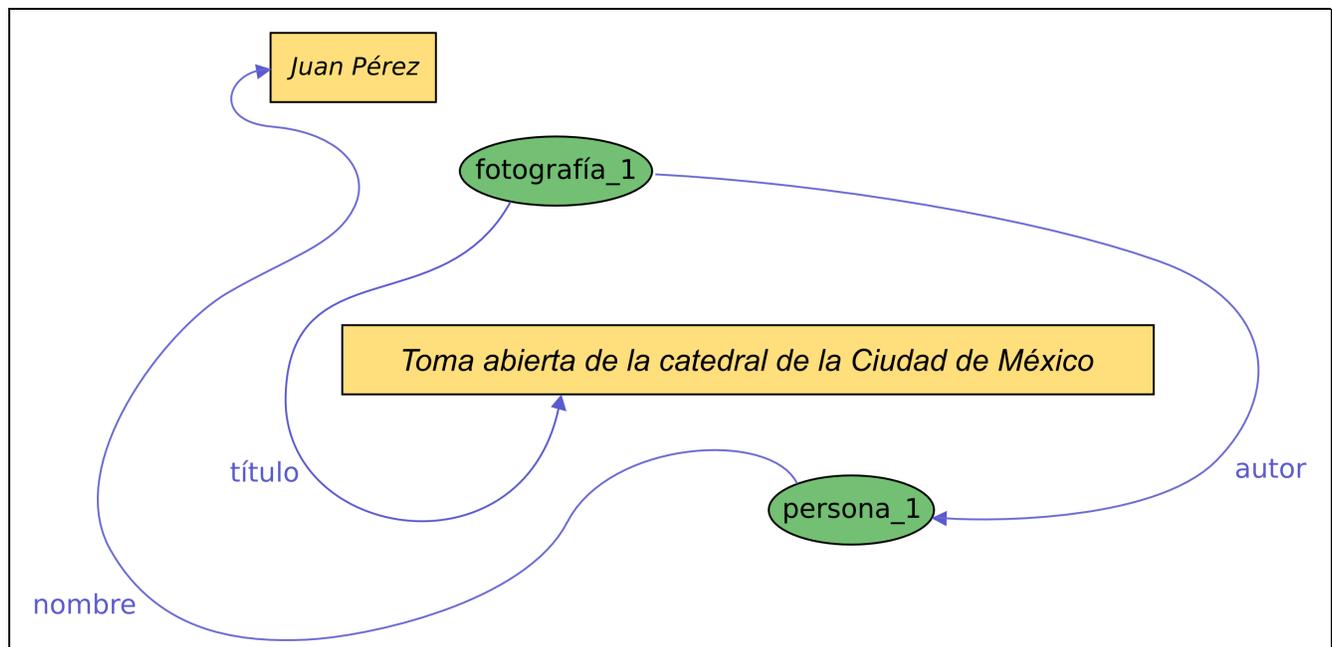
Título	Toma abierta de la catedral de la Ciudad de México
Autor	Juan Pérez

Figura 3. Otra reorganización de la misma información

Toma abierta de la catedral de la Ciudad de México / Juan Pérez.
--

Incluso podríamos concebir más formas de plasmar esta información, por ejemplo en un catálogo “hablante” u otro en braille. Lo que nos concierne aquí es la distinción que se establece entre la información que almacena y gestiona la RC, codificada como una red de relaciones entre significados, y los distintos formatos que se emplean para desplegar ante las personas dicha información. Desde la perspectiva de un catálogo que use esta tecnología, la versión de los datos contenida en la RC es más importante o más fundamental que las demás. Ahora, examinemos de nuevo la Figura 1, la cual, como mencioné, presenta el fragmento de la RC “tal cual”. Destaca que la información fundamental no es la gráfica en tanto *imagen* sino el conjunto de objetos y relaciones que esa gráfica describe. Por consiguiente, cualquier gráfica con los mismos elementos y conexiones entre ellos constituiría una representación igualmente fiel del mismo fragmento de RC—aunque fuera *visualmente* muy distinta. La Figura 4 es un ejemplo de una gráfica así.

Figura 4. Otra gráfica que representa el mismo fragmento de RC que la Figura 1



En otro texto expuse la tesis de que esta diferenciación entre una versión “fundamental” de los datos catalográficos (estructurados como una RC o bajo otro esquema similar) y otras versiones “de presentación” implica un cambio fundamental en cómo catalogamos. Recientemente han surgido varios sistemas y tecnologías que aceptan una separación de este tipo (ejemplos incluyen los Requisitos Funcionales para Registros Bibliográficos y las implementaciones de Dublin Core y MARC 21 en WS); en cambio dicha separación no forma parte de los sistemas basados en el papel o que trasladan directamente los mecanismos del papel al ámbito digital (como el MARC 21). Sostengo que los nuevos sistemas representan el inicio de un cambio de paradigma en las prácticas de catalogación, cuyo rasgo central es una separación entre los procesos de (1) codificación y estructuración de la información catalográfica, (2) comunicación de dicha información y (3) registro físico.⁶

Hace falta precisar algunos aspectos prácticos del uso de esta tecnología. Primero, la elaboración de los formatos de presentación de los datos a partir de los contenidos de una RC será un proceso automático; en absoluto se trata de llevar a cabo estas transformaciones manualmente. Sin embargo, aún está inmadura la tecnología que se requiere para éste y otros procesos que se llevarán a cabo en el marco del nuevo paradigma de la catalogación. La WS y otros tipos de RC no son “programas” con funciones sencillas y amigables, sino que son herramientas que los desarrolladores de *software* pueden integrar a sus programas y a partir de ellas crear funciones novedosas. Esto es lo que hace el sistema informático que desarrolla el proyecto PCHLA; integra la WS a sus mecanismo interno de procesamiento de información y con base en ella implementa determinadas funciones. Si bien ya se creó una primera versión de dicho sistema, ésta todavía no tiene funciones amigables para colocar en línea una colección digital, por lo que, por ahora, cualquier intento de generar un catálogo en WS por medio del sistema requiere de la intervención de un programador. Se espera que más adelante se reduzcan o se eliminen este tipo de barreras al uso de esta tecnología.

El empleo en la catalogación de las RC y metodologías relacionadas y la implantación del nuevo paradigma tienen numerosas ramificaciones. Por un lado se transforman los procesos de construcción de normas, ya que, para lograr el intercambio de datos entre distintos catálogos ya no será necesario buscar acuerdos acerca de los campos específicos que deben usarse para describir los materiales, sino

6 GREEN, Andrew R.: *Cambios actuales en el esquema de supuestos básicos de las prácticas catalográficas*. Ponencia presentada en el Seminario de Estudios sobre lo Visual en la Investigación Social, en el Instituto Mora, México D.F., el primero de septiembre de 2003. <http://www.nongnu.org/durito/docs/cambios_catalog_green.pdf>. En ese texto hablé de las RC como “estructuras abstractas” y sostuve:

[Bajo el nuevo paradigma] el primer paso en la generación de un catálogo será el análisis del conocimiento que se tiene acerca de los documentos y la elaboración de una estructura abstracta que permitirá representar este conocimiento. [...] Esta estructura abstracta será el componente central del catálogo. Se podrán elaborar múltiples expresiones del conocimiento que la estructura contendrá, en diferentes lenguajes y soportes físicos, para fines de comunicación entre personas o máquinas; pero el formato abstracto siempre será considerado el “original” o “molde” del cual las demás expresiones se derivarán.

consensuar los conceptos generales que organicen y atraviesen las RC. Esto se debe a la facilidad con que esta tecnología permite traducir entre distintos sistemas descriptivos—siempre que dichos sistemas empleen conceptos generales que son compatibles entre sí. (Véase por ejemplo el proyecto MarcOnt que emplea una RC para traducir entre los formatos MARC, Dublin Core y BibTeX.⁷) Ahora, destaca que las RC también pueden emplearse para codificar los sistemas de conceptos generales de este tipo; una RC de conceptos generales de una área de conocimiento determinado se llama “ontología de nivel medio”, y una RC de conceptos globales, cuya finalidad es construir puentes entre las ontologías de diferentes campos, es una “ontología de nivel alto”. A futuro, entonces, las normas de catalogación no serán hechas de reglas como “anote el año de publicación en el campo tal”, “use comillas si el título fue designado por el autor” o “escriba 's.l.' para materiales sin localizar”, sino que serán ontologías de nivel medio, compuestas por conceptos como “objeto patrimonial”, “preservación”, “manifestación” o “proceso de digitalización”, y reglas acerca de cómo esos conceptos deben vincularse entre sí y con las descripciones específicas de un catálogo. Asimismo los catálogos que se construirán con base en normas de este tipo incluirán una gran diversidad de formatos de ficha catalográfica y no obstante permitirán el intercambio de datos con mayor facilidad que los sistemas actuales.

Desde una perspectiva más general debemos integrar el análisis del empleo de las RC a las reflexiones teóricas y filosóficas que hacemos acerca del quehacer archivístico. Una línea de indagación muy prometedora es la que parte de la analogía entre las RC y los procesos de cognición humana. Resumidamente: las RC pueden verse como comparables, en algún sentido, con los modelos mentales a través de los cuales, según la ciencia cognitiva, las personas entendemos el mundo. Siguiendo esta analogía podemos entender las “versiones de presentación” de datos como similares al discurso, y las ontologías de nivel medio y alto se equipararían con la cultura de grupo y la cultura general respectivamente.⁸ Esta analogía permitiría construir un puente amplio entre diversas teorías sociales, nuestro entendimiento de la catalogación y la práctica de construir archivos. Con base en ella podríamos formular numerosas preguntas inéditas, por ejemplo: ¿qué implicaciones tienen para los archivos y sus catálogos las teorías de Habermas acerca de la acción comunicativa? o ¿cómo nos ayuda a construir conocimiento histórico a partir de los archivos de distintas culturas la noción de la “racionalidad cosmopolita” de Boaventura de Sousa?⁹

Finalmente, es importante construir un diálogo entre estas reflexiones y los debates actuales acerca de las bases filosóficas de la bibliotecología. Entre las muchas propuestas divergentes que han

7 Documentos acerca de ese proyecto pueden consultarse en el sitio <<http://portal.marcont.org/mediations.jsp>>.

8 Una de las numerosas teorías que podrían emplearse para desarrollar esta analogía es la que elabora Teun van Dijk acerca de la cognición, la ideología, el discurso, y la sociedad. VAN DIJK, Teun A.: *Ideología. Una aproximación multidisciplinaria*, Barcelona, España, 1999.

9 DE SOUSA SANTOS, Boaventura: “Hacia una sociología de las ausencias y una sociología de las emergencias”, en *El milenio huérfano. Ensayos para una nueva cultura política*, Madrid, Editorial Trotta, 2005, pp. 151-194.

surgido de este debate, destaca la de Luciano Floridi, quien argumenta que la bibliotecología debe abandonar sus sustento tradicional—la epistemología social, entendida como teoría del conocimiento social—y reemplazarlo con un nuevo campo denominado *filosofía de la información*. Su trabajo resulta especialmente relevante para el proyecto PCHLA en tanto la filosofía de la información originó en la filosofía de la inteligencia artificial, e incluso su definición como ámbito le permitiría abordar las ideas que acabo de presentar acerca de las RC.¹⁰

Conclusión

En esta ponencia describí de manera resumida el proyecto PCHLA y presenté algunas reflexiones que se generaron en su seno, principalmente acerca del marco interdisciplinario del proyecto y el uso de las RC en la catalogación. La amplitud de los temas abordados y de los problemas no resueltos nos señala que el generar archivos digitales para la investigación social es una empresa monumental; quizá se tarde varios años en analizar y resolver satisfactoriamente las dificultades centrales de la misma. Pero al mismo tiempo hay indicios del enorme potencial de la tecnología para transformar los procesos de generación de conocimiento. Fomentar esta transformación desde los principios de la ciencia social crítica y promoviendo el acceso libre y amplio al patrimonio cultural e histórico, así como al conocimiento que con base en él se genere, constituye una manera eficaz de valorar el patrimonio y llevar a la práctica dichos principios.

10 FLORIDI, Luciano: “On defining library and information science as applied philosophy of information”, en *Social Epistemology*, vol. 16 núm. 1 (2002), pp. 37-49.

ANEXO SISTEMA PESCADOR IMÁGENES DE PANTALLA Y FUNCIONES

El sistema Pescador es un *software* para servidor Web cuya función principal es generar las páginas de colecciones digitales en línea. La versión actual de este sistema se emplea en el sitio *Marcas de fuego de la Biblioteca “José María Lafragua”*.¹¹ Las opciones que presenta al usuario son todavía relativamente básicas: se puede visualizar el listado de marcas, revisar las descripciones detallada de las mismas, ver las fotografías en alta resolución y hacer búsquedas. Sin embargo en su funcionamiento interno el sistema es mucho más complejo de lo que aparenta: almacena el catálogo en una representación de conocimiento, de la cual extrae los datos “al vuelo” para construir las fichas que se soliciten. Se espera que este mecanismo permita a mediano y largo plazo la implementación de funciones especializadas para la consulta, análisis, gestión y catalogación de los fondos digitales.

La Figura 5 muestra una imagen de la pantalla con el listado de las marcas que registra el catálogo, descritas por medio de fichas abreviadas. Cuando el usuario hace *click* en el botón “Ver registro completo” se despliega la información detallada de la marca, tal como aparece en la Figura 6.

Destaca la organización de la ficha: la primera sección describe la marca en general, la segunda, la institución que la usó, y las demás, las manifestaciones específicas de la impronta y los libros en los que aparecen. Esta misma estructura queda reflejada en la RC, que hace un *modelaje de* los objetos que se describen.

Para visualizar las imágenes referidas en alta resolución el usuario simplemente hace *click* en cualquiera de las pestañas que aparecen arriba de la ficha.

Funciones proyectadas

Una meta central del proyecto PCHLA es ampliar el conjunto de funciones que proporciona el sistema Pescador. A corto plazo se implementarán la búsqueda avanzada y la navegación por facetas (que permitirá agrupar los elementos de una colección con base en los valores que presentan en determinados campos) así como mecanismos básicos para modificar tanto el contenido como la estructura de las fichas (éstos estarán disponibles únicamente para los encargados de los fondos digitales). A mediano y largo plazo se desarrollarán diversas funciones más avanzadas, las cuales se describen a continuación.

FUNCIONES CON METADATOS

Construcción y gestión de tesauros y estructuras similares

¹¹ El catálogo puede consultarse en <<http://www.marcasdefuego.buap.mx>>.

El sistema debe permitir la gestión de estructuras jerárquicas—incluyendo jerarquías con herencia múltiple y vínculos horizontales—para poder manejar adecuadamente objetos como tesauros, autoridades, temarios y taxonomías. Para ello, se debe implementar una interfaz Web que permita crear, modificar, vincular y navegar de manera sencilla y flexible estructuras de este tipo. Asimismo, es necesario implementar mecanismos de enlace automático con los servidores de taxonomías establecidas, e indagar mecanismos para “puentear” entre jerarquías diferentes que refieren un mismo tema.

Información geográfica

Los catálogos digitales suelen registrar la información geográfica por medio de campos de texto, pero un mecanismo más avanzado para ello integraría datos cuantitativos acerca de la ubicación y las geometrías de los lugares y regiones en la superficie del planeta. ¿Qué tipo estructuras en una RC, qué algoritmos para el procesamiento interno de datos y qué funciones para el usuario serían posibles y deseables para tratar la información geográfica? Esta pregunta debe tomarse como punto de partida para una investigación que dará lugar a sistemas de manejo y búsqueda más eficaces para los datos de este tipo.

Catálogos complementarios

En la catalogación a menudo es importante describir los documentos por medio de determinados términos y conceptos referentes al contexto social e histórico en que dichos materiales fueron producidos. Al llevar un control cuidadoso del uso de los elementos descriptivos de este tipo se incrementa la coherencia interna y utilidad a largo plazo de los catálogos generados. Por este motivo, el sistema debe incluir funciones para construir catálogos complementarios o glosarios (por ejemplo: de los tipos de vestimenta de una época y lugar determinado, o de los modelos y fabricantes de las máquinas empleadas por una industria específica) y permitir su vinculación con los elementos de un catálogo en línea.

Colecciones, conjuntos y funciones avanzadas de catalogación

El Pescador debe integrar diversas funciones para acelerar y consolidar los procesos de catalogación, entre ellas: la organización jerárquica de las fondos, colecciones y series y la definición de conjuntos de diversa índole; la asignación en bloque de metadatos a conjuntos de objetos; el desplazamiento de objetos de una colección o conjunto a otro; la vinculación de las diferentes versiones de una obra intelectual; la ingestión automática de catálogos y documentos digitales existentes; y un sistema para deshacer y rehacer cualquier cambio efectuado al archivo. Todas estas funciones deben ser accesibles desde la interfaz Web.

Exportación y conversión a normas

Resulta imprescindible que el sistema permita la exportación de los datos catalográficos en formatos

estándares como MARC 21 y Dublin Core. Asimismo debe poder generar diversos tipos de archivo (por ejemplo hojas de cálculo).

Formatos de documentos

Hasta ahora el desarrollo del Pescador se ha centrado en la difusión y estudio de archivos de imágenes, pero a futuro el sistema debe poder manejar con igual facilidad todo tipo de documento textual, de sonido y de imagen en movimiento.

ESPACIOS Y CONTRIBUCIONES DE LOS USUARIOS

Mecanismos que faciliten la investigación de los archivos

El Pescador debe facilitar la investigación de los archivos que difunde proporcionando a los usuarios mecanismos como el historial de búsqueda, canasta de objetos guardados, carpeta personal, y configuración personalizada del formato de los resultados de búsqueda, entre otros.

Catalogación continua

La catalogación de un archivo no debe terminar cuando el archivista lo pone a disposición de un público. El Pescador facilitará el mejoramiento constante de los datos catalográficos al permitir que los usuarios comenten la información que aparece en su pantalla, y que los comentarios estén disponibles en línea para los usuarios así como para los responsables de los fondos, quienes podrán evaluar las aportaciones así realizadas y actualizar los catálogos con base en un criterio profesional.

AMPLIANDO LA RED DE CONOCIMIENTO

Intercambios entre servidores

Existen protocolos establecidos para la difusión y recolección automatizada de los metadatos; el sistema Pescador debe poderlos aprovechar para aumentar el alcance de las búsqueda que realice. Asimismo, debe diseñarse un mecanismo de intercambio y búsqueda global en fondos almacenados en múltiples servidores que ejecuten el Pescador.

FUNCIONES ADICIONALES

Apariencia y documentos complementarios

El Pescador debe ofrecer una interfaz Web sencilla para configurar su apariencia y gestionar los documentos “complementarios”, y así poderse adaptar a las necesidades específicas de cada archivo. Es decir, proporcionará funciones similares a las que ofrecen los sistemas de gestión de contenido (CMS); sin embargo el Pescador, a diferencia de los CMS, las proporcionará en combinación con funciones

archivísticas.

Respaldos e informes

La gestión de fondos digitales requiere diversas funciones administrativas tales como el respaldo automático de los catálogos y documentos digitales, e informes estadísticos sobre los accesos al sitio Web y las búsquedas realizadas.

Figura 5 Listado de las marcas de fuego del catálogo

Marcas de fuego / Marcas de fuego

Archivo Editar Ver Web Ir Marcadores Solapas Ayuda

http://www.marcasdefuego.buap.mx/repository/list_view

Inicio Consultar el catálogo Acerca de las marcas de fuego

Marcas de fuego Todas las marcas del catálogo

Navegar 1 - 10 de 129 registros

Ordenar 0 seleccionados

Búsqueda

Hoja anterior Hoja siguiente

Reiniciar Ver

Buscar

Descripción
Forma
Identificador único

Identificador único: BJML/MF0101
Titular: Nuestra Señora de Gracia
Forma: A
Descripción: La Sigla "A" es la abreviatura de la palabra "Agustín", que sirvió para denotar que el libro perteneció a la Orden de San Agustín. Está trazada de forma irregular con dos fierros, uno largo casi triangular y otro más corto para marcar la barra de la A, ambos calentados al fuego casi al rojo vivo.. Esta marca está clasificada en las Marcas de Fuego Rudimentarias.
[Ver registro completo >>](#)

Identificador único: BJML/MF0102
Titular: Nuestra Señora de Gracia
Forma: A
Descripción: La Sigla "A" es la abreviatura de la palabra "Agustín", que sirvió para denotar que el libro perteneció a la Orden de San Agustín. Está trazada de forma irregular con dos fierros, uno largo casi triangular y otro más corto para marcar la barra de la A, ambos calentados al fuego casi al rojo vivo. Difiere de la marca BJML/MF0101 en que las puntas de la barra son hacia abajo.
[Ver registro completo >>](#)

Identificador único: BJML/MF0103
Titular: Nuestra Señora de Gracia
Forma: A
Descripción: La Sigla "A" es la abreviatura de la palabra "Agustín", que sirvió para denotar que el libro perteneció a la Orden de San Agustín.
[Ver registro completo >>](#)

Identificador único: BJML/MF0104
Titular: Nuestra Señora de Gracia
Forma: A
Descripción: La Sigla "A" es la abreviatura de la palabra "Agustín", que sirvió para denotar que el libro perteneció a la

http://www.marcasdefuego.buap.mx/repository/list_view?page_num=0&order_by=vw_hasUniquelidentifier&list_uri=rep_soc_00...

Figura 6 Ficha detallada de una marca de fuego

Marcas de fuego / Marca de fuego

Inicio Consultar el catálogo Acerca de las marcas de fuego

Registro completo

Regresar Navegar Búsqueda

Registro 20 de 129

Registro anterior Registro siguiente

Buscar

Ficha completa Imagen Portada del libro Anotación manuscrita

Marca de fuego	
Identificador único	BJML/MF0301
Clase	Heráldico
Categoría	Corporativo
Forma	Cuadrangular redondeado
Descripción	Escudo cuadrangular de base redondeada, sin esmaltes: un monte hueco sumado de una cruz, cargado de una letra capital N y acostado de las letras capitales A y G (N y G en minúsculas, que en conjunto forman la abreviatura ANG de ANGELES), al timbre, corona abierta.
Referencia	Sala, Rafael. Marcas de fuego de las antiguas bibliotecas mexicanas. México: Monografías Bibliográficas Mexicanas / SRE, 1925. Marca 123, pág. 73.
Referencia	Krausse Rodríguez, Carlos. Marcas de fuego. Catálogo. México: Biblioteca Nacional de Antropología e Historia, 1989. Pág. 42.
Observaciones	En el libro de las Religiosas Carmelitas Descalzas del convento de "San José y Santa Teresa" de esta ciudad de Puebla, encontramos la siguiente descripción heráldica de la Orden Carmelita Descalza: "Campo de plata con monte de sable, con tres estrellas, dos en jefe, sable sobre plata y una en punta plata sobre sable. Al escudo está sobre puesta una corona ducal". Rosalba Loreto López (Editora). Una empresa Divina. Las hijas de Santa Teresa de Jesús en América. 1604- 2004. Puebla, Consejo Cultural UDLA, Puebla, 2004, p. XVI.

Titular	
Tipo de institución	Convento Masculino
Advocación original	Nuestra Señora de los Remedios
Nombre ordinario	Nuestra Señora del Carmen
Ubicación	Ciudad de Puebla, Puebla
Orden	Orden de los Carmelitas Descalzos
Provincia	San Alberto de México

Ejemplar	
Imagen	
Dimensiones	35 x 25 mm
Ubicación de la marca	Canto superior
Libro	De Jesús María, Juan. Cantici canticorum interpretario / Per F. Ioannem à Iesu Maria, Carmelitan Discalceatum, Calaguritanum. Salamaticæ: Ex officina Art. Tabernelij Antuerpiani, Excudebat Artus Tabernellius Antuerpianus, 1602. Catálogo de la Biblioteca >>
Portada del libro	
Anotación manuscrita	
Transcripción de la anotación	"del convento de Ntra. Señora de los Remedios de la Puebla de Carmelitas descalzos"
Ubicación de anotación	Portada
Evidencias de propiedad histórica	Se estableció la propiedad del documento con base en la anotación manuscrita.

Ejemplar	
Libro	Días, Philippe, (O.F.M.). Marial de la Sacratissima Virgen nuestra señora: en que se contienen muchas consideraciones de grande spiritu y puntos delicatissimos de la diuina scriptura [...]. Salamanca: Juan Fernandez, 1596. Catálogo de la Biblioteca >>
Transcripción de la anotación	"Del Convento de Nuestra Señora de los Remedios de Carmelitas Descalzos"
Ubicación de anotación	Portada