

ANÁLISIS Y CATEGORIZACIÓN DE LOS DATOS ABIERTOS DE LAS BIBLIOTECAS MUNICIPALES ESPAÑOLAS: METADATOS, INTEROPERABILIDAD Y PROPUESTA PARA LA APERTURA Y REUTILIZACIÓN

Analysis and categorization of the open data of Spanish municipal libraries: metadata, interoperability, and proposal for the opening and reuse

Diego Maseda-Seco, Gema Bueno-de-la-Fuente y Eva Méndez



Diego Maseda-Seco es webmaster en la *Rede de Bibliotecas Municipais da Coruña*, responsable de la producción, el mantenimiento y la actualización de los contenidos de la web y del software de gestión de actividades e inscripciones. Trabaja en la automatización de la obtención de datos e indicadores de la colección mediante *Microstrategy*. Ha colaborado en el desarrollo de una taxonomía maestra transversal municipal basada en la norma técnica de interoperabilidad de reutilización de recursos de información. Master en *Bibliotecas y servicios de información digital* por la *Universidad Carlos III de Madrid*, licenciado en Documentación por la *Universidade da Coruña* y en Humanidades por la *Universidade de Santiago de Compostela*.
<http://orcid.org/0000-0002-2423-0615>

Bibliotecas Municipais da Coruña, Servizo Municipal de Bibliotecas
Rúa Durán Loriga, 10 - 4ª. 15003 A Coruña, España
d.maseda@coruna.es



Gema Bueno-de-la-Fuente es profesora ayudante doctor del *Departamento de Ciencias de la Documentación* de la *Universidad de Zaragoza*, y anteriormente en el *Departamento de Biblioteconomía y Documentación* de la *Universidad Carlos III de Madrid*. Durante el curso 2015/16 trabajó como *EU Projects Officer* en la *European Association of Research Libraries (Liber)* con sede en La Haya. Doctora en Documentación, sus líneas de trabajo se centran en datos abiertos, datos de investigación, bibliotecas y repositorios digitales, vocabularios y web semántica, y preservación digital.
<http://orcid.org/0000-0001-5400-1488>

Universidad de Zaragoza, Departamento de Ciencias de la Documentación
Pedro Cerbuna, 12. 50009 Zaragoza, España.
gbueno@unizar.es



Eva Méndez es profesora titular del *Departamento de Biblioteconomía y Documentación* de la *Universidad Carlos III de Madrid*, donde es vicerrectora adjunta de estrategia y educación digital. Doctora en documentación, su docencia e investigación se centra en metadatos, web semántica, infraestructuras de información digital, acceso abierto, políticas de información y web social. Pertenece al *Advisory Board* de *Dublin Core (DCMI)* y al de *OpenAire*. Es miembro del comité ejecutivo de *Rebiun* y de la *European Open Science Policy Platform* de la *Comisión Europea*.
<http://orcid.org/0000-0002-5337-4722>

Universidad Carlos III de Madrid, Facultad de Humanidades, Comunicación y Documentación
C/ Madrid, 128. 28903 Getafe (Madrid), España.
emendez@bib.uc3m.es

Resumen

Se analiza el estado de la apertura de los datos de bibliotecas municipales en España como parte constitutiva de la información del sector público. El artículo estudia información de los portales de administraciones municipales españolas que sirven conjuntos de datos abiertos de sus bibliotecas. Se categorizan dichos conjuntos de datos y se analizan, a modo de ejemplo, las propiedades, atributos o metadatos utilizados por tres bibliotecas que exponen conjuntos de datos sobre actividades (A Coruña, Madrid y Las Palmas) determinando su potencial interoperabilidad desde el mapeo de vocabularios preexistentes aplicables. Asimismo, el artículo recoge una propuesta para priorizar la publicación de conjuntos de datos potencialmente más interoperables, así como una breve revisión inicial del contexto normativo y legal para la apertura de datos del sector público en España.

Artículo recibido el 10-01-2017

Aceptación definitiva: 08-05-2017

Palabras clave

Datos abiertos; Conjuntos de datos; Reutilización de información; Sector público; Bibliotecas públicas; Bibliotecas municipales; Interoperabilidad; España.

Abstract

The current situation of open data at local public libraries in Spain is analyzed as a part of public sector information. This article extracts information from data portals at the city level that include datasets produced by public libraries. These datasets are categorized by type of data (offer/demand of services and infrastructures). Three libraries that expose datasets about activities (A Coruña, Madrid and Las Palmas) are selected in order to compare properties, attributes, and metadata used to describe these datasets, analyzing their potential interoperability from the mapping of applicable existing vocabularies. This article also proposes a prioritization of the publication of more interoperable datasets and a brief initial review of the legal and regulatory context for the opening of public sector data in Spain.

Keywords

Open data; Information reuse; Public sector; Public libraries; Municipal libraries; Interoperability; Spain.

Maseda-Seco, Diego; Bueno-de-la-Fuente, Gema; Méndez, Eva (2017). "Análisis y categorización de los datos abiertos de las bibliotecas municipales españolas: metadatos, interoperabilidad y propuesta para la apertura y reutilización". *El profesional de la información*, v. 26, n. 3, pp. 392-402.

<https://doi.org/10.3145/epi.2017.may.05>

1. Introducción y contexto: reutilización de la información del sector público

La reutilización de la información del sector público se refiere al uso que hacen ciudadanos y empresas de las informaciones que la administración pública genera en el ejercicio de sus funciones. Implica también una redistribución de la información en tanto que la administración libera sus datos para mejorar su eficiencia y permitir el crecimiento económico del sector privado, que a su vez revertirá en el incremento del bienestar social.

La *Directiva 2013/37/UE del Parlamento Europeo y del Consejo*, de 26 de junio de 2013, por la que se modifica la *Directiva 2003/98/CE* relativa a la reutilización de la información del sector público (*Unión Europea*, 2013), reconoce el valor económico de la publicación de datos abiertos, pero también su valor social por su potencial para el aumento de la transparencia y la exigencia de responsabilidades y fiscalización de la labor de los poderes públicos. La información que se genera y se gestiona en las administraciones públicas es (o debería ser) una información de calidad, completa y fiable que, si se publica y reutiliza, permite un aprovechamiento óptimo para la generación de nuevos, y probablemente innovadores, productos, servicios y mercados por parte de terceros. La administración pública, por su parte, debe asumir un proceso de apertura de datos para facilitar el intercambio y el análisis de la información pública para adaptarla a cualquier fin (Eaves, 2009).

Fruto de procesos *open data* o datos abiertos, los ciudadanos y las organizaciones pueden acceder libremente a estos datos para informarse o para crear nuevos servicios y aumentar el

valor social y, si es el caso, también el valor comercial de esta información (Garriga-Portolà, 2011). Además, esta apertura y compartición aumenta la transparencia administrativa y la ética, a las que se une también, según Garriga-Portolà (2011), la idea de la democratización del acceso a la información pública y la eliminación de élites privilegiadamente informadas. Estas ideas son compatibles con los principios y la razón de ser de las bibliotecas públicas y municipales.

Ahora bien, no toda la información pública disponible o publicada en la Web es información abierta válida para su reutilización. No sólo se trata de publicar los datos sino que hay que garantizar el acceso a ellos, razón por la que debe recurrirse a formatos digitales, estandarizados y abiertos, siguiendo una estructura clara que permita su comprensión y reutilización (Garriga-Portolà, 2011). Es clarificadora la conocida y asimilada clasificación de Berners-Lee (2006) en relación con los datos abiertos, que utiliza estrellas, como los hoteles, para determinar el nivel de apertura y potencial reutilización, explotación y consumo de los datos (figura 1).



Figura 1. Esquema de niveles de 5 estrellas para datos abiertos (Hausenblas, 2012)

1.1. Marco regulador

La *Directiva* europea de 2013 (*Unión Europea*, 2013) insta a las administraciones a abrir sus datos, en tanto que se trata de una información que puede desempeñar:

“una función importante a la hora de impulsar el desarrollo de nuevos servicios basados en formas novedosas de combinar y utilizar esa información, estimular el crecimiento económico y promover el compromiso social”.

Por su parte, los fondos contenidos en bibliotecas, museos y archivos:

“constituyen una base potencial para productos y servicios de contenidos digitales y tienen un enorme potencial de reutilización innovadora en sectores tales como el aprendizaje y el turismo”.

En nuestro país, el *Real decreto 1495/2011*, de 24 de octubre, que desarrolla la *Ley 37/2007*, de 16 de noviembre, sobre reutilización de la información del sector público (*España*, 2011), reconoce la importancia y el valor de la información que se genera en las administraciones públicas por lo mucho que puede influir sobre la economía y el crecimiento del empleo. Con el objetivo de guiar la acción de las administraciones públicas, en 2009 el *Ministerio de Industria, Turismo y Comercio*, a través de la *Secretaría de Estado de Telecomunicaciones y para la Sociedad de la Información*, publicó la *Guía Aporta sobre reutilización de la información del sector público (Plan Avanza2, 2009)*, documento que recoge los aspectos clave a tener en cuenta para dar el impulso definitivo a la apertura de los datos y que constituye un fuerte apoyo a la difusión y aplicación de la normativa española en torno a la reutilización.

La democratización del acceso a la información pública y la eliminación de élites privilegiadamente informadas son dos ideas perfectamente compatibles con los principios y razón de ser de las bibliotecas públicas y municipales

Hasta la entrada en vigor de la *Ley 18/2015* de 9 de julio, que modifica la *Ley 37/2007 (España, 2015)*, las bibliotecas se encontraban fuera del alcance de dicha normativa. Ahora:

“se ha ampliado el ámbito de aplicación a las bibliotecas, incluidas las universitarias, los museos y los archivos, dado el importante volumen de recursos de información que poseen y los proyectos de digitalización que vienen llevando a cabo”.

1.2. Agentes reutilizadores de los datos abiertos

Se pueden identificar tres agentes o roles en el proceso de reutilización de la información del sector público (*Plan Avanza2, 2009*):

1) Administración pública: juega un papel central en la reutilización en tanto que es quien recoge, genera y reproduce una gran variedad y cantidad de información en el ejercicio de sus funciones. Su papel es crucial y su finalidad doble:

- de cara a las empresas, persigue contribuir al desarrollo económico de éstas y a la creación de empleo;
- frente a los ciudadanos, consigue aumentar su transparencia y fomentar la participación.

2) Entidades infomediarias: interesadas en cualquiera de los tipos informativos que emanan de la Administración, utilizan esta información para generar productos o servicios con valor añadido llevando a cabo también nuevos usos en la industria de los contenidos digitales. La creación de nuevas empresas y puestos de trabajo son consecuencia directa de estos nuevos productos y servicios.

3) Usuarios finales: ciudadanos y empresas, que son los destinatarios finales de la información liberada y reutilizada.

En el ejercicio de sus funciones, el sector público produce una muy variada y valiosa información susceptible de ser reutilizada por la industria de contenidos digitales o por la propia ciudadanía

1.3. Valor económico de la reutilización de la información

En un solo minuto en internet, las administraciones públicas, las empresas y los ciudadanos en general producen un vertiginoso cúmulo de datos. Por ejemplo, se lanzan más de cuatro millones de búsquedas en *Google*, se publican más de 347.000 tweets, se suben más de 100 horas de vídeo a *YouTube*, se descargan más de 194.000 aplicaciones o se suben 38.000 fotos a *Instagram (OBS Business School, 2015)*.

El sector público produce en el ejercicio de sus funciones una variada y valiosa información susceptible de ser reutilizada por la industria de contenidos digitales o por los ciudadanos. Información geográfica, meteorológica, cultural, turística o estadística, son sólo algunos ejemplos que constituyen una materia prima perfecta para nuevos productos de información digital. El estudio *Measuring European public sector information resources (Dekkers et al., 2006)*, estima el valor del mercado de la reutilización de la información pública entre los 26 y los 47 mil millones de euros en la Unión Europea. Esta cifra en 2016 asciende a más de 55 mil millones de euros y para el intervalo 2016-2020 se pronostica que el tamaño del mercado directo acumulado para los datos abiertos será de 325 mil millones de euros y que se crearán 25.000 nuevos puestos de trabajo (*European Commission, 2016*).

1.4. Evolución de la publicación de información pública en España

La tendencia de publicar datos en abierto va en aumento. El portal *datos.gob.es* nació en 2009 fruto del *Proyecto Aporta*, con el fin último de organizar y gestionar los catálogos de información del sector público. En marzo de 2014 *datos.gob.es* tenía registrados 1.609 conjuntos de datos provenientes de las administraciones públicas de España (**Martínez-Méndez; López-Carreño; Pastor-Sánchez, 2014**). En septiembre de

2016 el buscador del portal ofrecía un total de 11.817 resultados. En la figura 2 se muestra la evolución del número de conjuntos de datos asociados a alguna de las categorías temáticas en ambos años (un mismo conjunto de datos puede tener asociadas una o más categorías).

En 2015 la *Asociación Española de Normalización y Certificación* publicó la norma técnica en español sobre *smart cities: Ciudades inteligentes. Datos abiertos* (Aenor, 2015) con el objetivo de servir de guía para las administraciones y organismos públicos en la definición, documentación e implementación de proyectos de datos abiertos. Entre otras cosas, establece indicadores y métricas de uso para medir los grados de madurez de la apertura de datos de una ciudad y poder realizar comparaciones. **Khokhar** (2016) en su estudio de los portales de datos abiertos de los países europeos otorga a *datos.gob.es* el grado de "pionero". En el análisis tiene en cuenta aspectos como la reutilización de los datos que contienen o la usabilidad del portal, y establece tres niveles:

- principiantes: países que cuentan con portales de datos abiertos muy básicos en cuanto a funciones y cobertura de conjuntos de datos (como República Checa, Lituania o Letonia);
- seguidores: países con portales más avanzados, pero con apertura aún limitada (como Polonia, Rumanía o los Países Bajos);
- pioneros: los países europeos con portales de datos abiertos más avanzados, con funciones más sólidas (como España, Reino Unido o Francia).

Aunque el portal *datos.gob.es* nació para aglutinar y dar acceso a todos los datos públicos abiertos, hay administraciones que cuentan con un portal propio y que no realizan un volcado de sus conjuntos de datos al portal nacional. Esto sucede por ejemplo en ayuntamientos como los de A Coruña, Sevilla o Barcelona entre otros, lo cual dificulta la utilización de *datos.gob.es* como única fuente de análisis para este estudio.

1.5. Bibliotecas, apertura de datos y reutilización de información

Casi de manera generalizada las bibliotecas recaban multitud de datos referidos a usos, inversiones, actividades, colección, solicitudes, etc., que facilitan la toma de decisiones, el análisis de resultados o simplemente, la realización de un seguimiento de los servicios y su evolución. Aunque las bibliotecas constituyen un gran potencial de datos de calidad y fiables (...), hasta ahora sólo los han producido y compartido en su propio medio, con poco contacto con otros colectivos (**Peset; Ferrer-Sapena; Subirats-Coll**, 2011).

Cualquiera puede hacer un subproducto con los datos, y las bibliotecas deben facilitar que otros agentes reutilicen sus datos para darles una nueva vida y un nuevo ciclo donde

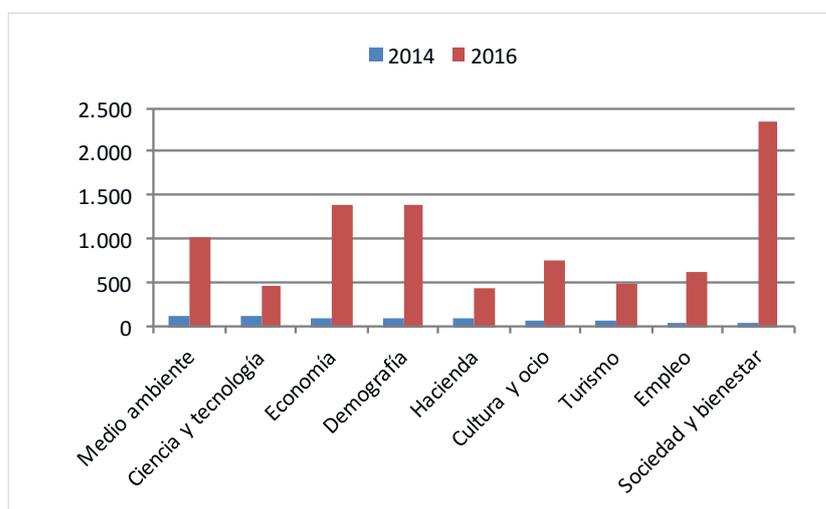


Figura 2. Conjuntos de datos publicados en 2014 y 2016 en *datos.gob.es*, por tipo de datos

otros puedan explotarlos. Es momento de que las bibliotecas se sitúen en el centro del movimiento de datos abiertos y se constituyan en infomediarias (**Saorín**, 2011). Con este objetivo la *Biblioteca Nacional de España* publicó en marzo de 2016 su *Plan de medidas de impulso de la reutilización de la información* (BNE, 2016).

Las bibliotecas públicas y las municipales son el máximo exponente de instituciones culturales del sector público ya que pertenecen a la Administración y, por ende, generan información y datos de carácter público. Así por ejemplo, deberían liberar o publicar como datos abiertos sus catálogos, entre otros, los metadatos que generan en sus catálogos. Por todo ello en el apartado siguiente nos centraremos en analizar qué datos publican actualmente nuestras bibliotecas municipales, con el interés de reflejar el *statu quo* de la apertura de este tipo de datos de la administración local, así como detectar y analizar su prospectiva de reutilización e interoperabilidad.

En principio, cualquiera puede obtener subproductos de los datos, y las bibliotecas deben facilitar que otros agentes reutilicen sus datos para darles una nueva vida, un nuevo valor, y un nuevo ciclo donde otros puedan explotarlos

2. Análisis de la apertura de datos en bibliotecas municipales españolas

2.1. Objetivos

El objeto de estudio son las bibliotecas municipales como un elemento fundamental de la administración municipal en la apertura de datos, así como un componente de las ciudades inteligentes.

El objetivo general es determinar el papel que las bibliotecas municipales españolas están jugando en el entorno de los datos abiertos dentro de las principales administraciones locales del país. Los objetivos específicos son:

- identificar los municipios españoles que publican datos abiertos de o sobre sus bibliotecas municipales;
- analizar la oferta y la demanda de información que puede y debe servirse como datos abiertos;
- clasificar los principales conjuntos de datos abiertos de las bibliotecas municipales;
- analizar las principales propiedades o elementos de metadatos que describen los conjuntos de datos detectados para revisar su potencial interoperabilidad y reutilización;
- realizar una reflexión o recomendación para el proceso de publicación efectiva de datos de bibliotecas.

2.2. Metodología

Para conocer los datos abiertos de y sobre bibliotecas municipales en España y elaborar una propuesta de proceso o *workflow* para la publicación de conjuntos de datos, se han seguido los siguientes pasos:

- identificación de los portales de datos municipales que publican conjuntos de datos de y sobre bibliotecas;
- selección y recolección de conjuntos de datos publicados de y sobre bibliotecas municipales;
- categorización de los conjuntos de datos según el tipo de datos que proporcionan: oferta y demanda de contenidos y servicios, y directorios de instalaciones;
- análisis comparativo de propiedades o elementos de metadatos que describen los conjuntos de datos de un tipo seleccionado;
- mapeo de los elementos identificados con otros esquemas de metadatos o vocabularios existentes para la descripción de entidades de similares características.

Para identificar los portales de datos de alcance municipal, se partió del mapa de iniciativas de *datos.gob.es*, a partir de los portales de "Administración local". Del total de 61 portales incluidos, se descartaron los de diputaciones o cabildos, multi-ayuntamientos y los específicos de datos de transpor-

te, obteniéndose un listado final de 49. Se realizaron búsquedas en cada uno para localizar conjuntos de datos de o sobre bibliotecas municipales, siguiendo distintas estrategias dependiendo del número de conjuntos de datos, opciones de búsqueda y navegación del portal. En su mayoría se pueden localizar clasificados en la categoría "Cultura y ocio" combinada con la palabra clave "bibliotecas". El 48,9% (24 de los 49 portales seleccionados) publican o distribuyen al menos un conjunto de datos sobre bibliotecas municipales o publicados por ellas.

“Casi un 50% de los portales de la administración municipal seleccionados para este estudio publican o distribuyen al menos un conjunto de datos sobre bibliotecas municipales o publicados por ellas”

Una vez identificados los portales se extrajeron los datos de cada uno, obteniendo un total de 45 conjuntos de datos, que se categorizaron en función del tipo de información que ofrecen.

Para analizar los elementos de metadatos (propiedades o atributos) más comúnmente utilizados y analizar su potencial interoperabilidad, se trató de identificar los ayuntamientos con información más homogénea que permitieran comparar y reflejar sus posibilidades de reutilización¹.

2.3. Resultados

2.3.1. Análisis y categorización de los conjuntos de datos

Al analizar el contenido de los 45 conjuntos de datos de los 24 portales que finalmente se consideraron para el estudio, se encontraron tres categorías en función del tipo de información que ofrecen sobre: oferta y demanda de contenidos

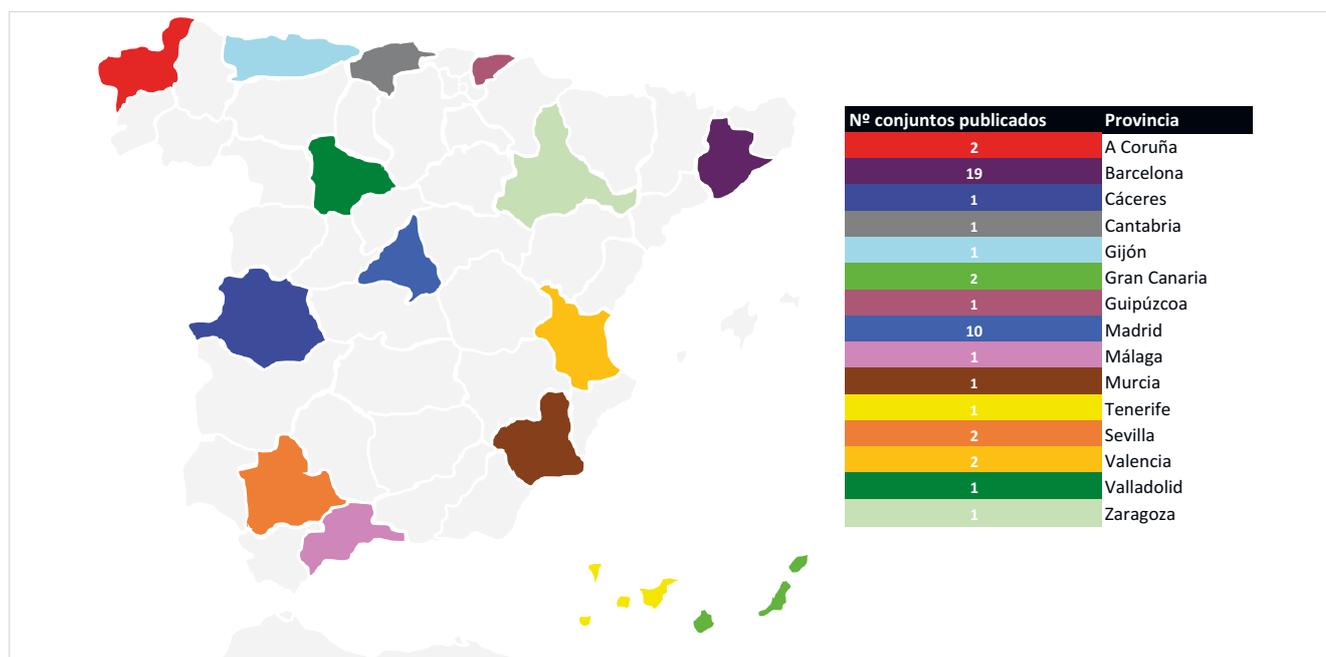


Figura 3. Provincias con municipios que publican datos de bibliotecas y número de conjuntos de datos publicados

y/o servicios; instalaciones y equipamientos municipales (tabla 1) (en algunos casos un conjunto de datos se ha clasificado bajo dos categorías):

- oferta: se reúnen 15 conjuntos de datos sobre servicios

o prestaciones que ofrecen las bibliotecas a la sociedad (ejemplos: fondos bibliográficos disponibles para su uso, o actividades que se realizan en estas instalaciones);

- demanda: se incluyen 12 conjuntos de datos con informa-

Tabla 1. Datos abiertos de las bibliotecas municipales españolas por municipio y categoría de información

Ayuntamiento	Oferta	Demanda	Directorios de instalaciones
A Coruña	Actividades de las bibliotecas		Bibliotecas municipales
Alcobendas		Mediatecas municipales. Datos de usuarios y tipos de uso	
Arganda del Rey	Biblioteca digital (catálogo)		
	Hemeroteca (catálogo)		
	Mediateca (catálogo)		
Arona			Instalaciones culturales
Barcelona	Fondo documental de las bibliotecas	Usuarios de los archivos y bibliotecas patrimoniales	Listado de equipamientos de cultura y ocio de la ciudad de Barcelona
		Consultas en los archivos y bibliotecas patrimoniales	
		Visitas a las bibliotecas	
		Asistentes a las actividades de difusión cultural y de lectura de las bibliotecas	
		Asistentes a las visitas escolares a las bibliotecas de Barcelona	
		Préstamos realizados en las bibliotecas	
		Usos de ordenadores y de wifi en las bibliotecas de Barcelona	
Cáceres			Bibliotecas Cáceres
Camargo			Bibliotecas en el municipio de Camargo
Cornellà de Llobregat	Estadísticas de bibliotecas	Estadísticas de bibliotecas	
Gavà			Listado de equipamientos de la ciudad
			Mapa de equipamientos de la ciudad
Gijón			Bibliotecas
Granollers	Indicadors. Cultura. Biblioteques	Indicadors. Cultura. Biblioteques	
Las Palmas de Gran Canaria	Crono_eventos (Eventos culturales)		Crono_espacios (espacios culturales)
Lorca			Bibliotecas en el municipio de Lorca
Madrid	Actividades gratuitas en Bibliotecas Municipales en los próximos 60 días	Bibliotecas Municipales: Préstamos activos	Bibliotecas y bibliobuses de la ciudad de Madrid
	Bibliotecas Municipales: Catálogo	Bibliotecas Municipales: préstamos activos. Histórico	Sedes: bibliotecas especializadas universitarias y nacionales
Málaga			Telecentros
Manlleu	Fons documental <i>Biblioteca Mossèn Blancafort</i>		
	Fons documental <i>Biblioteca Bisbe Morgades</i>		
Sabadell	Activitats Biblioteques		
	Dades de biblioteques		
San Sebastián	Noticias de bibliotecas (canal RSS)		
Sevilla			Bibliotecas
			Equipamientos municipales
Terrassa			Equipamientos de la ciudad
Torrent			Entitats culturals i museus
Valencia			Equipamientos municipales
Valladolid			Relación de bibliotecas de uso público en Valladolid
Zaragoza			Equipamientos

ción sobre los usos que los ciudadanos hacen de las bibliotecas o que demandan de ellas (ejemplos: conjuntos de datos de préstamos, consultas o uso de ordenadores y wifi);

- directorios de instalaciones municipales: 20 conjuntos de datos sobre la localización física y las características de los equipamientos municipales de las administraciones (ejemplos: bibliotecas y bibliobuses de Madrid, el listado de equipamientos de cultura y ocio de la ciudad de Barcelona).

La mayor parte de los portales de datos abiertos municipales que se han analizado publican datos recogidos sobre bibliotecas en la categoría de directorios de instalaciones municipales, en algunos casos, con datos referidos exclusivamente a bibliotecas, o agrupadas en mediatecas y espacios culturales.

La segunda categoría más común sería la de oferta de contenidos y servicios, si bien en varios casos se trata de datos estadísticos o indicadores que proporcionan un bajo nivel de detalle sobre dicha oferta.

Los datos sobre demanda de servicios es la categoría menos común. Únicamente las bibliotecas de Madrid, Barcelona, Granollers, Cornellà de Ll. y Alcobendas, proporcionan algún dato relativo a los usuarios, préstamos, uso de ordenadores y asistencia a actividades.

El interés de suministrar en formatos abiertos la información de los fondos documentales de las bibliotecas es ampliamente reconocido no sólo en el contexto de la información pública sino también, y especialmente, por su potencial en el ámbito de los datos enlazados (Byrne; Goddard, 2010; Baker et al., 2012; Saorín, 2012).

La información sobre ubicación geográfica de los equipamientos municipales suele ser compartida también por las administraciones autonómicas reuniendo en un único conjunto de datos toda la información de los equipamientos municipales de una provincia.

La información estadística es una información con potencial que se recaba también de forma normalizada por el Consejo de Cooperación Bibliotecaria y resulta de indudable interés para las bibliotecas y para sistemas de bibliotecas nacionales, autonómicos y municipales, pero a priori no presenta potencialidades de explotación para el usuario final.

En cambio, la información sobre actividades o eventos organizados por las bibliotecas es la menos analizada hasta la fecha y, quizá debido a ello, la menos frecuente. Sólo las administraciones locales de A Coruña, Madrid, Las Palmas y Sabadell están compartiendo conjuntos de datos referidos a ello. Nos centraremos en las tres primeras localidades, en

Tabla 2. Comparativa de los principales atributos/metadatos utilizados para la descripción de conjuntos de datos de "Actividades" en A Coruña, Madrid y Las Palmas

A Coruña		Madrid		Las Palmas	
Descripción de la propiedad	Elemento de metadato	Descripción de la propiedad	Elemento de metadato	Descripción de la propiedad	Elemento de metadato
Campo de uso interno en el gestor de contenidos, que se asigna automáticamente	Nombre	Campo clave de uso interno	ID-evento	Campo clave de uso interno	Id
Campo de descripción del tipo de actividad: por tipo de público, por tipo de actividad, por tema y por ubicación	Tipo	Campo de descripción del tipo de actividad	Tipo	Campo de descripción de la tipología de la actividad	Categoría
Campo que recoge el título del evento a que se refiere	Título	Campo que recoge el título del evento a que se refiere	Título	Campo que recoge el título del evento a que se refiere	Nombre
Campo de fecha de inicio del evento, con la estructura AAAA-MM-DD HH:MM:SS.0	Fecha inicio	Campo de fecha de inicio del evento, con la estructura AAAA-MM-DD HH:MM:SS.0	Fecha	Campo de fecha de inicio del evento, con la estructura AAAA-MM-DD HH:MM:SS.0	Fecha inicio
Campo de fecha de fin del evento, con la estructura AAAA-MM-DD HH:MM:SS.0	Fecha fin	Campo de fecha de fin del evento, con la estructura AAAA-MM-DD HH:MM:SS.0	Fecha fin	Campo de fecha de fin del evento, con la estructura AAAA-MM-DD HH:MM:SS.0	Fecha fin
Campo de descripción del evento	Cuerpo	Campo de descripción del evento	Descripción	Campo de descripción del evento	Descripción
Campo donde se indica la hora de inicio del evento, con la estructura HH:MM	Horarios	Campo donde se indica la hora de inicio del evento, con la estructura HH:MM	Hora		
Campo donde se especifica la instalación donde se lleva a cabo el evento	Lugar de celebración	Campo donde se especifica la instalación donde se lleva a cabo el evento	Nombre instalación	Campo donde se especifica la instalación donde se lleva a cabo el evento	Ubicación
Campo de coordenada de 6 caracteres sobre plano de la ciudad	Coordenada X	Campo de coordenada de 6 caracteres sobre plano de la ciudad	Coordenada X		
Campo de coordenada de 7 caracteres sobre plano de la ciudad	Coordenada Y	Campo de coordenada de 7 caracteres sobre plano de la ciudad	Coordenada Y		
Campo donde se vincula el evento con el programa al que pertenece	Contenido relacionado	Campo donde se vincula el evento con el programa al que pertenece	Título actividad		

tanto que comparten información de actividades que se celebran únicamente en las bibliotecas. Sabadell, por su parte, ofrece conjuntos de datos con información de actividades organizadas por todos los servicios del municipio, no sólo de bibliotecas.

Algunas de las propiedades podrían mapearse con propiedades y atributos de otros vocabularios estándar para la descripción de eventos de una agenda cultural, como *schema.org* o *Event*

2.3.2. Análisis de elementos de metadatos y mapeo a otros vocabularios

Los conjuntos de datos de la categoría “oferta” son todos de carácter estadístico, referidos a datos totales de número de documentos, préstamos, visitas, usuarios activos de préstamo, adquisiciones. Pero sólo los conjuntos de datos relativos a “Actividades” de los ayuntamientos de A Coruña, Madrid y Las Palmas describen el contenido u otras características de las actividades en forma de propiedades relativamente normalizadas a través de elementos de metadatos. En la tabla 2 se muestra una comparación de estos elementos.

Con el objetivo de equiparar las propiedades que describen a nivel interno los conjuntos de datos en las citadas administraciones, se comparan con los principales esquemas semánticos que se utilizan para describir conjuntos de datos semejantes a los de este análisis. Algunas de las propiedades podrían mapearse con propiedades y atributos de otros vocabularios estándar para la descripción de eventos de una agenda cultural, como *schema.org* o *Event ontology* (voca-

bulario para la representación de eventos en una agenda cultural) (Corcho, 2015) que, a su vez reutiliza conceptos de otros vocabularios o modelos que se están usando con propósitos similares.

En el vocabulario *Event ontology* (figura 4), la clase “Evento” cuenta con los siguientes atributos o propiedades:

- Accesible
- Fecha de fin
- Fecha de inicio
- Hora de fin
- Hora de inicio
- Lugar de inscripción
- Medio de transporte
- Servicio municipal
- Tipo de accesibilidad
- Tipo de evento.

Si mapeamos los atributos o elementos de este vocabulario con los utilizados en la descripción del conjunto de datos “Actividades”, obtenemos el resultado que se muestra en la tabla 3.

El tema de los datos abiertos está de moda, pero aún es muy nuevo y cuenta con pocas experiencias en bibliotecas municipales

En definitiva, se ha comprobado que existe una alta correspondencia entre los elementos de ambos conjuntos de datos, así como con los del vocabulario escogido. Se considera factible agregar conjuntos de datos de estas características e incluso normalizar y mapear los datos con vocabularios

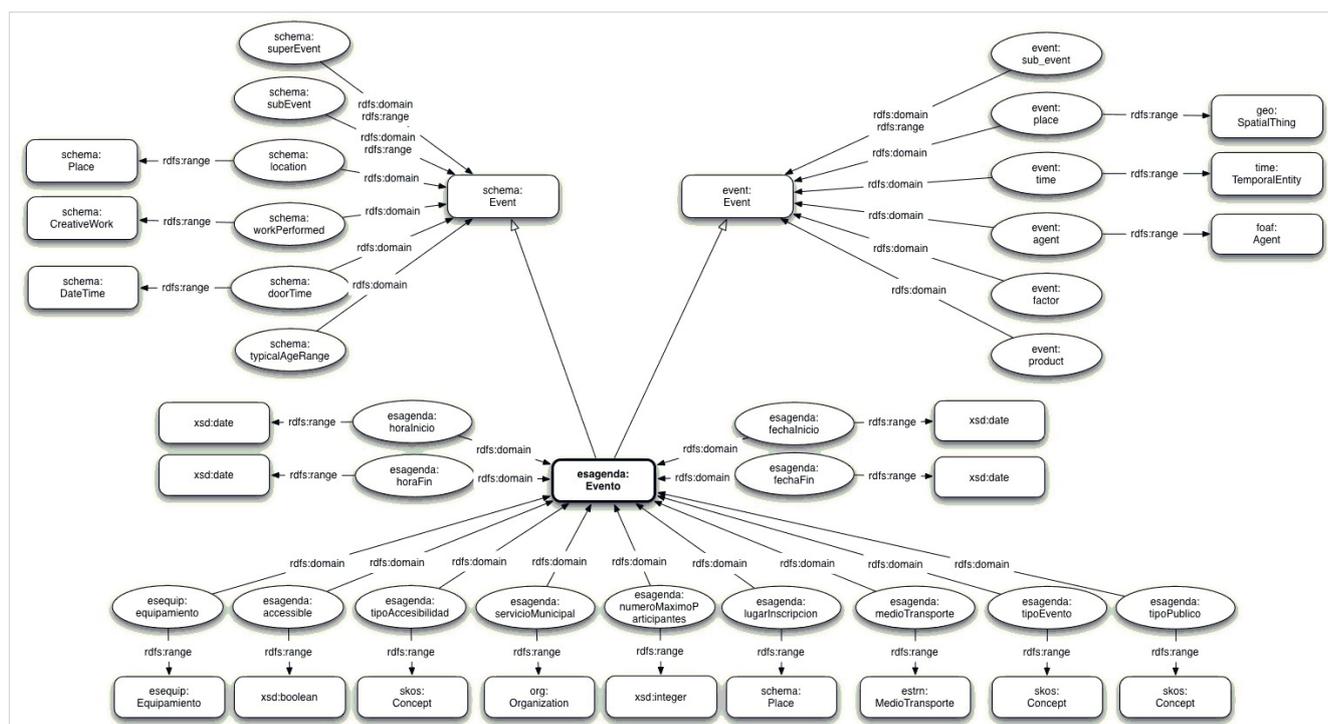


Figura 4. Definición de la clase “Evento” (Corcho, 2015)

de uso compartido con otros ámbitos e instituciones para facilitar su reutilización y potenciar la elaboración de nuevos productos y servicios que tengan como materia prima los datos de bibliotecas.

2.4. Propuesta de flujo de trabajo para la apertura de conjuntos de datos en bibliotecas municipales

El objetivo final de cualquier iniciativa de datos abiertos en bibliotecas sería la publicación de todos los datos disponibles y aptos para ser abiertos. Sin embargo, abrir todos los datos puede ser complejo y costoso; sobre todo teniendo en cuenta la escasez generalizada de recursos y limitaciones de tiempo que sufren estas bibliotecas. Además, el asunto

de los datos abiertos está de moda pero aún es muy nuevo y cuenta con pocas experiencias en bibliotecas municipales, como hemos visto. Por todo ello y teniendo en cuenta los resultados de este análisis, recomendamos comenzar por abrir o exponer aquellos conjuntos de datos más estandarizados y potencialmente interoperables desde el punto de vista de los metadatos que utilizan en sus descripciones.

Se propone un flujo de trabajo de cinco fases o pasos a la hora de publicar conjuntos de datos de bibliotecas municipales (figura 5).

Paso 1: Fase de análisis

Se divide en dos sub-fases: análisis externo e interno.

En primer lugar se realizará un análisis del entorno donde se revisarán trabajos previos de apertura de datos en bibliotecas: qué se abre y cómo se expone (qué metadatos se utilizan).

Posteriormente se realizará un análisis interno o diagnóstico que permita dar respuesta a cuestiones como: ¿qué datos estoy almacenando?, ¿puede mi organización afrontar el proceso de publicación?, ¿las herramientas donde almacenamos los datos a nivel interno, permiten una extracción sencilla?, etc. En resumen, se trata de identificar las mejores prácticas para adaptarlas y aplicarlas a nivel interno.

Tabla 3. Mapeo de los elementos que describen los conjunto de datos “Actividades” con los elementos del vocabulario para la representación de eventos

Elementos en el conjunto de datos “Actividades”	Elementos mapeables al vocabulario “Eventos” URI base: <i>http://vocab.linkeddata.es/datosabiertos/def/cultura-ocio/agenda</i>
Nombre	
Tipo En A Coruña se divide en: tipo de actividad, tipo de público, tema y ubicación	#tipoEvento
	#tipoPublico
Título	#Evento
Fecha inicio	#fechalnicio
Fecha fin	#fechaFin
Cuerpo	
Horarios	#horaFin
	#horalnicio
Lugar de celebración	#equipamiento
Coordenada X	
Coordenada Y	
Contenido relacionado	
Precio	

Paso 2: Fase de selección y mapeo de datos

Realizado el inventario de datos internos, conviene determinar cuáles son más valiosos desde el punto de vista de un agente reutilizador. También hay que descartar aquellos conjuntos de datos sobre los que operen restricciones de carácter legal como la protección de datos de carácter personal, la seguridad pública o el derecho de propiedad intelectual, etc. Habrá que determinar sobre qué conjuntos se han de realizar operaciones de filtrado y sobre cuáles no, etc. Se trata de priorizar la publicación de datos en función de su potencial audiencia y de la complejidad de su publicación. Durante el mapeo se tratará de vincular los datos de la biblioteca con los de otras organizaciones. Para ello es fundamental identificar elementos de metadatos y propiedades o atributos compartidos con otros. Si además de mapear propiedades podemos enlazar o interconectar datos, se estará potenciando su reutilización.

Paso 3: Fase de filtrado y extracción

Es imprescindible analizar con detenimiento los conjuntos de datos y metadatos, propiedades y atributos que los describen y someterlos a los procesos de filtrado necesarios para respetar las normas legales de la publicación de datos abiertos (ejemplo: procesos de anonimización para eliminar toda referencia a la identidad de las personas). Para la ex-



Figura 5. Flujo de trabajo para la toma de decisiones de apertura de datos en bibliotecas municipales

tracción de datos a partir de la fuente donde se almacenan es imprescindible automatizar el proceso de tal forma que los datos sean fácilmente extraíbles desde la biblioteca municipal donde se encuentren, al portal del ayuntamiento o al portal de la administración donde se quieran publicar o compartir.

Paso 4: Fase de preparación, descripción y asignación de licencias

Preparar los datos para su publicación final implica, por un lado, completar los valores de los metadatos que describen el conjunto de datos para fomentar la recuperación y, por tanto, la potencial reutilización de los datos. Por otra parte es necesario asignar las licencias necesarias que permitan a cualquier agente reutilizador tener en cuenta las condiciones de uso que se hayan marcado sobre los datos. Tanto a la hora de completar los registros de metadatos como de asignar licencias a los datos de las bibliotecas municipales, debemos pensar en los potenciales reutilizadores que podrían sacar provecho de los conjuntos de datos en cuestión.

Paso 5: Fase de publicación y compartición de datos

Nutrir el catálogo de datos con los conjuntos de datos procedentes de bibliotecas municipales, publicitarlos y compartirlos en catálogos o portales de datos para darles mayor visibilidad será el fin último de este flujo de trabajo (figura 5).

En esta última fase se visibilizarán los datos de las bibliotecas, así como sus elementos, propiedades y valores descriptivos, potenciando su reutilización. Como se ha anticipado, la fase de publicación es el resultado final, pero se fundamenta en que la extracción sea automática y en la normalización de su descripción.

Contar con un flujo de datos concreto y perfectamente definido beneficiará la inclusión de conjuntos de datos procedentes de bibliotecas municipales

3. Conclusiones

A la vista del análisis sobre conjuntos de datos publicados por las administraciones municipales sobre bibliotecas, y del análisis pormenorizado de las propiedades y atributos que los describen, se pueden extraer las siguientes conclusiones de este trabajo:

- La publicación de información abierta por parte de bibliotecas municipales en España es todavía incipiente: sólo se encuentran 45 conjuntos de datos. De ellos, aún son pocos los que cuentan con una descripción normalizada de las propiedades de los conjuntos de datos mediante metadatos.
- La mayor parte de las bibliotecas municipales activas en portales de datos abiertos proporcionan datos referidos a la geolocalización de sus instalaciones. En segundo lugar se publican conjuntos de datos en la categoría de oferta de contenidos y servicios, y en menor medida sobre demanda de servicios, con datos sobre usuarios, préstamos, uso de ordenadores y asistencia a actividades.

- Es necesario priorizar la apertura de los datos internos que tienen su equivalencia en los vocabularios estándar y, por tanto, son comúnmente aceptados y utilizados, de tal forma que se potencie desde el principio su integración e interoperabilidad.
- Contar con un *workflow* concreto y perfectamente definido beneficiará la inclusión de conjuntos de datos procedentes de bibliotecas municipales, tanto en portales generales de la Administración, como en colecciones de datos disponibles desde portales de los ayuntamientos.

Nota

1. El análisis de los 24 portales seleccionados para el estudio se ha publicado en *Figshare* como hoja de cálculo mostrando los datos crudos y la comparativa del estudio. Los datos pueden encontrarse en: *Analysis of datasets published as opendata of municipal libraries in Spain*. Figshare. <https://doi.org/10.6084/m9.figshare.4967801.v2>

4. Referencias

Aenor (2015). *Norma UNE 178307:2015. Ciudades inteligentes: datos abiertos (open data)*. Asociación Española de Normalización y Certificación, 21 enero.

Baker, Thomas; Bermès, Emmanuelle; Coyle, Karen; Dunsire, Gordon; Isaac, Antoine; Murray, Peter; Panzer, Michael; Schneider, Jodi; Singer, Ross; Summers, Ed; Waites, William; Young, Jeff; Zeng, Marcia (2012). *Library Linked Data Incubator Group. Final report*. W3C Incubator Group. <http://www.w3.org/2005/Incubator/llid/XGR-llid-20111025>

Berners-Lee, Tim (2006). *Linked data*. 27 July. <https://www.w3.org/DesignIssues/LinkedData.html>

Biblioteca Nacional de España (2016). *Plan de medidas de impulso de la reutilización de la información*. Marzo. http://www.bne.es/webdocs/LaBNE/Transparencia/Institucional/plan_risp_bne.pdf

Byrne, Gillian; Goddard, Lisa (2010). "The strongest link: Libraries and linked data". *DLib magazine*, v. 16, n. 11/12. <http://www.dlib.org/dlib/november10/byrne/11byrne.html>

Corcho, Óscar (2015). *Vocabulario para la representación de eventos de una agenda cultural*, 6 marzo. <http://vocab.linkeddata.es/datosabiertos/def/cultura-ocio/agenda-content/0.1/index.html>

Dekkers, Makx; Polman, Femke; Te-Velde, Robbin; DeVries, Marc (2006). *Mepsir - Measuring European public sector information resources. Final report of study on exploitation of public sector information- benchmarking of EU framework conditions*. http://ec.europa.eu/newsroom/document.cfm?doc_id=1197

Eaves, David (2009). "Las tres leyes de los datos abiertos". *Eaves.ca*, 11 diciembre. <https://eaves.ca/2009/12/11/las-tres-leyes-de-los-datos-abiertos>

España (2011). "Real decreto 1495/2011, de 24 de octubre, por el que se desarrolla la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público, para el ámbito del sector público estatal". *BOE*, n. 269, 8 noviembre.

https://www.boe.es/diario_boe/txt.php?id=BOE-A-2011-17560

España (2015). "Ley 18/2015, de 9 de julio, por la que se modifica la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público". *BOE*, n. 164, 10 julio.

https://www.boe.es/diario_boe/txt.php?id=BOE-A-2015-7731

European Commission (2015). *Creating value through open data*. European Data Portal. ISBN: 978 9279527913

https://www.europeandataportal.eu/sites/default/files/edp_creating_value_through_open_data_0.pdf

Garriga-Portolà, Marc (2011). "¿Datos abiertos? Sí, pero de forma sostenible". *El profesional de la información*, v. 20, n. 3, pp. 298-303.

<https://doi.org/10.3145/epi.2011.may.08>

Hausenblas, Michael (2012). *5 estrellas datos abiertos*.

<http://5stardata.info/es>

Khokhar, Tariq (2016). "European countries making clear progress with open data". *The World Bank. The data blog*, 30 March.

http://blogs.worldbank.org/opendata/european-countries-making-clear-progress-open-data?CID=ECR_TT_worldbank_EN_EXT

Martínez-Méndez, Francisco-Javier; López-Carreño, Rosana; Pastor-Sánchez, Juan-Antonio (2014). "Open data en las administraciones públicas españolas: categorías temáticas y apps". *El profesional de la información*, v. 23, n. 4, pp. 415-424.

<https://doi.org/10.3145/epi.2014.jul.10>

OBS Business School (2015). "En 2020, más de 30 mil millones de dispositivos estarán conectados a Internet". *OBS Business School*.

<http://www.obs-edu.com/noticias/estudio-obs/en-2020-mas-de-30-mil-millones-de-dispositivos-estaran-conectados-internet>

Peset, Fernanda; Ferrer-Sapena, Antonia; Subirats-Coll, Inma (2011). "Open data y linked open data: su impacto en el área de bibliotecas y documentación". *El profesional de la información*, v. 20, n. 3, pp. 165-173.

<https://doi.org/10.3145/epi.2011.mar.06>

Plan Avanza2 (2009). *Guía Aporta sobre reutilización de información del sector público*. Madrid: Plan Avanza2; MiTyC; Red.es; Aporta.

<http://www.datos.gob.es/content/guia-aporta>

Saorín-Pérez, Tomás (2012). "Cómo linked open data impactará en las bibliotecas a través de la innovación abierta". *Anuario ThinkEPI*, v. 6, pp. 288-292.

<http://www.thinkepi.net/como-linked-open-data-impactara-en-las-bibliotecas-a-traves-de-la-innovacion-abierta>

Unión Europea (2013). "Directiva 2013/37/UE del Parlamento Europeo y del Consejo, de 26 de junio de 2013, por la que se modifica la Directiva 2003/98/CE relativa a la reutilización de la información del sector público". *Diario oficial de la Unión Europea*, n. 175, 27 de junio.

https://www.boe.es/diario_boe/txt.php?id=DOUE-L-2013-81251

Da visibilidad a tu trabajo depositándolo en e-LIS, el mayor repositorio internacional sobre biblioteconomía, documentación y comunicación

e-LIS e-prints in library & information science

Search | Home | Login | Create Account

Welcome to e-LIS Repository

Latest additions

Search the repository using a full range of fields

<http://eprints.rclis.org>