

CAPÍTULO 1.

INTRODUCCIÓN

Gema Bueno-de-la-Fuente (Univ. Carlos III)
David Rodríguez-Mateos (Univ. Carlos III)

Versión preprint de la obra publicada:

Bueno de la Fuente, Gema; Rodríguez Mateos, David. Introducción. En: Bueno de la Fuente, Gema (coord.) *La Iniciativa de Archivos Abiertos (OAI): situación y perspectivas en España y Latinoamérica*. Bogotá: Rojas Eberhard, 2007, pp. 1-15. ISBN 978-958-9121-89-4.

Sitio web del editor: http://www.rojaseberhard.com.co/rojaseberhard/bibliotecologia/lib_oai.html.

1.1. INTERNET Y LA DIFUSIÓN DE INFORMACIÓN CIENTÍFICA

Los centros de investigación y las universidades de todo el mundo se han visto inmersos en el auge de Internet como un movimiento que se convertía, en poco más de década y media, en una poderosa herramienta de comunicación, también para la información académica y científica. Hoy en día, es técnicamente posible acceder a casi cualquier contenido digital publicado en cualquier lugar del mundo. Todo científico puede, al menos desde el punto de vista técnico, hacer públicos los resultados de sus investigaciones y consultar los equivalentes de los colegas en sus áreas de investigación, de forma generalmente inmediata.

Esta afirmación, sin embargo, dista mucho de ser totalmente cierta. Existen todavía numerosas barreras, ya sean técnicas, económicas o humanas, que hacen que este acceso a la información electrónica sea aún difícil, cuando no imposible: por la mera falta de conexión a Internet para buena parte de la población mundial; porque la distribución de contenidos científicos a través de canales comerciales plantea problemas económicos incluso a las universidades y centros de investigación más importantes del mundo, o porque la diseminación de los contenidos a través de numerosos canales hace mucho más difícil su localización a través de Internet.

Para los profesionales de la información, ya sea en su rol de académicos (profesores o investigadores en las diversas universidades y centros de investigación de todo el mundo), o de bibliotecarios o especialistas de los centros de documentación, Internet les plantea, por lo tanto, un enorme reto: resolver el problema del acceso a cualquier documento científico de interés, y hacerlo de la manera más fácil, más eficiente y más eficaz posible.

En paralelo con este interés, numerosos investigadores comenzaron ya a mediados de los 90 a clamar contra esta paradoja: en la Sociedad de la Información, en la que los desarrollos tecnológicos posibilitan la publicación y el acceso en el medio digital, continúa siendo difícil acceder a los resultados de las investigaciones de otros colegas. Y en muchos casos, el autor ve mermado el acceso a sus propios documentos, restringidos por las publicaciones científicas que permiten sólo su consulta a quienes pueden pagar unas cuotas cada vez más caras y restrictivas. Unas publicaciones que, por otra parte, eran y son uno de los principales puntos de referencia para medir las carreras profesionales de los científicos y profesores de universidad en todo el mundo.

1.2. LA BÚSQUEDA DE ALTERNATIVAS DE ACCESO A LOS DOCUMENTOS CIENTÍFICOS

Mucha más información, más fácilmente accesible, más necesaria en unos tiempos de continuo cambio y de rápido avance de la ciencia, pero más difícil de obtener. Estos fueron los principales argumentos que propiciaron, que los investigadores de algunas disciplinas punteras (como la física, la economía o la psicología), más acostumbrados a compartir información por canales informales, fomentaran la creación de depósitos en red en los que difundir sus publicaciones, antes o después de ser incluidas en sus respectivas revistas de prestigio.

Esos depósitos fueron los archivos de *e-prints* (o versiones electrónicas de un artículo científico), posteriormente llamados repositorios temáticos, que fueron surgiendo progresivamente desde finales de los 90 en diversas áreas del saber. Poco a poco, viendo el éxito inicial en las áreas temáticas donde ya existían archivos (como los pioneros Arxiv en ciencias físicas, o RePEC en economía), las que hasta el momento eran iniciativas aisladas de investigadores e instituciones concretas comenzaron a materializarse en proyectos de envergadura. Aumentaron así los fondos de los investigadores que aceptaban ser incluidos en esos repositorios, se elevó el número de los investigadores interesados en participar en ellos y fue ampliándose el número de áreas científicas en las que surgían archivos de *e-prints*.

Pero hacía falta algo más para que dejaran de ser iniciativas aisladas en un ámbito, el de la comunicación científica, aún dominado por la mera difusión comercial. Dos iniciativas, distintas pero muy relacionadas, surgieron para dar forma a un movimiento cada vez más extendido: la primera, de tipo conceptual, se conoce como el movimiento de acceso abierto a los contenidos científicos, denominado abreviadamente *acceso abierto* o, en inglés, *open access* (OA).

Esta iniciativa debía ser complementada con otra herramienta más técnica. Los distintos repositorios que iban surgiendo funcionaban con diversos programas diferentes entre sí, lo cual hacía difícil el tratamiento de información proveniente de diferentes repositorios. Hacía falta un mecanismo que permitiera lograr la interoperabilidad entre todos ellos; en otras palabras, que hiciera posible que, al menos las descripciones de cada uno de los documentos de un repositorio pudieran ser expresadas en un lenguaje común, de forma que fueran fácilmente compartidas y recolectadas por otros programas.

Con ello, los datos de los diferentes repositorios podrían ser incluidos en nuevas herramientas de búsqueda que localizaran esos contenidos y ofrecieran una mínima descripción, facilitando su búsqueda y su recuperación. El movimiento que impulsa este propósito es la llamada Iniciativa de Archivos Abiertos (OAI).

El principal logro de la OAI ha sido la creación de un protocolo o lenguaje de comunicación entre las diferentes herramientas o sistemas que comparten información. Con ello, los datos de los diferentes repositorios podrían ser incluidos en nuevas herramientas de búsqueda que localizaran esos contenidos y ofrecieran una mínima descripción, facilitando su búsqueda y su recuperación. Este protocolo se denomina OAI-PMH (Open Archives Initiative – Protocol for Metadata Harvesting).

1.3. FUNDAMENTO DE LA OBRA

El protocolo OAI-PMH, desde su publicación en 2001, ha propiciado la aparición de toda una serie de herramientas imprescindibles no solo para que los profesionales de la información científica puedan recuperar los contenidos generados en otras instituciones y por otros científicos, sino también para que los fondos de sus propios investigadores sean puestos a disposición de la comunidad científica nacional e internacional.

Este libro pretende mostrar un panorama sobre el funcionamiento de OAI-PMH y de los diferentes productos basados en el protocolo: cuáles fueron sus orígenes; cuáles son sus fundamentos técnicos; qué productos basados en OAI-PMH pueden ser utilizados para la creación y difusión de contenidos científicos; cuál es la situación de estos productos en todo el mundo y, especialmente, en nuestras áreas de referencia: en Latinoamérica, por una parte, y en España, por otra; cuáles son las principales herramientas de software empleadas para todo ello y, por último, recopilar una lista de recursos donde obtener más información sobre un tema que está en plena expansión.

Los contenidos han sido elaborados por una serie de investigadores que han participado en proyectos relacionados con la implantación y el desarrollo de servicios basados en OAI-PMH en España y en Colombia desde 2005, tanto en el ámbito universitario (Universidad Carlos III de Madrid y Universidad de Salamanca, en España; y Universidad de La Salle, en Colombia) como en el ámbito privado (en la empresa española Digibis).

Uno de los objetivos declarados de este texto es animar a los profesionales de la información científica y académica, en el área hispana, para que se planteen la creación de servicios y herramientas interoperables conforme al protocolo OAI-PMH, que aumenten la difusión y faciliten el acceso a la información científica, basándose en las experiencias previas tanto en su área geográfica como en el resto del mundo.

No obstante, la presente obra no pretende ser un manual sobre la creación de estos servicios (obras que han comenzado a publicarse ya aunque hasta el momento sólo en inglés), sino una primera introducción al tema de carácter divulgativo. Si bien se abordan varios aspectos más técnicos (como en el capítulo 3, dedicado a explicar la especificación del protocolo OAI-PMH, o en el capítulo 8, que desglosa las diferentes herramientas de software que pueden emplearse), no se realiza una descripción en profundidad de los mismos, ni de los requisitos informáticos necesarios para su puesta en marcha.

Se trata, por lo tanto, de una obra pionera en el ámbito hispano, recomendable para aquellos bibliotecarios o documentalistas que no conozcan en profundidad el protocolo OAI-PMH ni sus múltiples posibilidades, así como para cualquier otro investigador que esté interesado en la difusión de información científica y en su puesta a disposición de otros investigadores, preferiblemente, siguiendo la filosofía del acceso abierto.

1.4. PARTES DE LA OBRA

La obra, además de la presente introducción, está dividida en ocho capítulos más, así como otra serie de recursos complementarios, como un glosario y un índice de siglas.

El capítulo 2 pretende describir las diferencias y relaciones entre el concepto de acceso abierto y la Iniciativa de Accesos Abiertos, así como el origen y las implicaciones que conllevan llevan ambos movimientos.

Los dos capítulos siguientes se centran en el protocolo OAI-PMH desde un punto de vista técnico: describiendo la especificación concreta del protocolo y las diferentes partes del mismo (capítulo 3) para, a continuación, detallar las diferentes herramientas que utilizan OAI-PMH para la difusión y recuperación de información científica (capítulo 4).

Los capítulos 5, 6 y 7 tienen un carácter más descriptivo: pretenden reflejar la situación actual de la difusión de OAI-PMH, mediante una panorámica internacional (capítulo 5), y haciendo luego hincapié en dos áreas concretas: Latinoamérica (capítulo 6) y España (capítulo 7).

Por último, el capítulo 8 presenta las diferentes aplicaciones y programas de software empleados en todo el mundo para la creación de proveedores de servicios y proveedores de datos que emplean el protocolo, mientras que el capítulo 9 es una selección de recursos sobre el tema.

1.4.1. Diferencias y relaciones entre *acceso abierto* y la Iniciativa de Archivos Abiertos

El capítulo 2, “El movimiento Open Access y la Open Archives Initiative”, explica el concepto de acceso abierto, así como la diferencia entre éste y el de repositorios o archivos abiertos.

El acceso abierto, en palabras de Hernández, consiste en “el acceso libre (en el sentido de gratuito) a la documentación científica, facilitando su lectura, copia, uso y distribución, aún manteniendo el reconocimiento del copyright de los artículos (y otras publicaciones como informes, libros, etc.) a sus autores”. Por otra parte, un repositorio o archivo abierto se entiende como “interoperable en cuanto al uso de una tecnología específica que permita que el contenido albergado por un archivo (o más exactamente un repositorio) sea accesible para motores de búsqueda web, para recolectores de metadatos, etc.”.

Este capítulo explica también el contexto en el que surge el acceso abierto, comenzando por sus motivaciones y las primeras declaraciones teóricas al respecto, elaboradas en diversas reuniones científicas celebradas en las ciudades de Budapest (Hungría), Berlín (Alemania) y Bethesda (Estados Unidos) y de las cuales toman el nombre. Estas declaraciones son consideradas el origen oficial del movimiento, si bien éste ya había surgido de modo espontáneo a mediados de los 90. En los últimos años, otras declaraciones o manifestaciones de apoyo al movimiento por el acceso abierto a la

ciencia han sido promulgadas por diversas instituciones oficiales, universitarias y científicas de todo el mundo.

Por último se describen las dos vías principales de difusión científica mediante acceso abierto: la publicación en una revista que permita el acceso abierto (vía dorada), o bien, la publicación en una revista que no sea de acceso abierto, pero auto-archivando el contenido, además, en un repositorio abierto (vía verde).

1.4.2. OAI-PMH: explicación y desarrollo

El resto de la obra está específicamente dedicada a OAI-PMH y a los desarrollos basados en el protocolo. El capítulo 3, "El protocolo OAI-PMH", que resulta el más complejo por su carácter eminentemente técnico, describe el protocolo OAI-PMH en profundidad, en su actual versión 2.0.

Tras una breve introducción histórica sobre sus orígenes, se describen los conceptos básicos del protocolo, que permite la descripción de recursos (objetos digitales o no digitales) mediante diversos formatos de descripción (o esquemas de metadatos) y regula el intercambio de dichas descripciones entre sistemas de información. OAI-PMH está basado en tecnologías ya existentes como: el lenguaje XML, para la elaboración de los documentos OAI-PMH; el formato de metadatos básicos de Dublin Core, para la definición de los metadatos posibles para la descripción del contenido; y el protocolo HTTP, para la conexión entre los diversos servicios que usan OAI-PMH.

El capítulo 4, "Implementación de OAI-PMH", se centra en las herramientas que, basándose en OAI-PMH, ofrecen, localizan o recopilan contenidos de tipo científico en cualquier lugar de Internet. En síntesis, se describe el papel que, en el marco del protocolo OAI-PMH, juegan los *proveedores de datos*, o depósitos que almacenan las descripciones en OAI-PMH de los diversos objetos; los *proveedores de servicios*, o herramientas que recolectan esos metadatos de diversos proveedores de datos para incluirlos en una base común (incluyendo el identificador de cada ítem, que permite acceder a él) y facilitan una la búsqueda y el acceso integrado a recursos científicos provenientes de diversas fuentes; y por último, de un híbrido entre ambos, la figura de los *agregadores*.

El capítulo comienza con algunas consideraciones generales que han de tenerse en cuenta en el momento de implementar un proveedor de datos o un proveedor de servicios: los objetivos y las políticas de uso que se pretenden, los metadatos que se podrán emplear y en qué formato, si se emplearán categorías temáticas, qué herramientas de software utilizar, etc.

A continuación, se explican en detalle las características de los proveedores de datos y los proveedores de servicios. Se detallan sus requisitos técnicos y funcionales, así como su arquitectura. También se muestra la posibilidad de realizar tests de prueba para comprobar que funcionan correctamente, así como, en el caso de los proveedores de datos, la recomendación de inscribir el repositorio generado, ya sea de tipo temático o de tipo institucional, en alguno de los directorios ya existentes, con el fin de dar a conocer su existencia.

Asimismo, se explica el concepto de agregador, o herramienta híbrida entre un proveedor de servicios (porque recolecta datos de uno o varios proveedores de datos) y proveedor de datos (porque los contenidos recogidos pueden ser de nuevo recolectados por otros proveedores de servicios)., y se describen dos figuras especiales: los repositorios estáticos (generados a partir de los contenidos almacenados en una base de datos, pero convertidos en un único documento XML fijo) y las pasarelas que hacen posible el acceso a estos repositorios estáticos.

1.4.3. Los proveedores de datos

El capítulo 5, "Aplicaciones y difusión del protocolo OAI-PMH", pretende ofrecer una visión global de la implantación de OAI en el mundo. Se toman como fuente de datos dos de las principales recopilaciones sobre repositorios: el registro de archivos abiertos ROAR y el directorio de repositorios de acceso abierto OpenDOAR. No obstante, la recogida de datos globales resulta incompleta, teniendo en cuenta que la existencia de muchos repositorios no se difunde suficientemente, y que el número total recogido por las distintas fuentes oscila entre los 900 repositorios de ROAR y OpenDOAR y los 1500 repositorios del registro de la Universidad de Illinois en Urbana Champaign),

De los datos de ambos servicios se extrae que los países con mayor número de repositorios conforme al protocolo OAI-PMH son Estados Unidos, Alemania y el Reino Unido. Otros países como Australia, Holanda, Canadá, Francia, Brasil o España aparecen en los primeros lugares, aunque a una considerable distancia.

Por materias, la mitad de los repositorios analizados por los distintos registros son multidisciplinares mientras que, de los repositorios centrados en materias concretas, destacan los de medicina y ciencias de la salud, informática y telecomunicaciones, ciencia en general e historia y arqueología, aunque existen repositorios en otras veinte disciplinas distintas. La mayoría de los repositorios aún no han definido políticas de uso de sus metadatos.

En cuanto a los contenidos, algo más de la mitad de los repositorios son institucionales, destacando las tesis como tipo de contenido más abundante, aunque también se incluyen informes y documentos de trabajo, preprints y postprints, así como comunicaciones en congresos científicos, entre otros.

Junto a esta visión global, el capítulo incluye una amplia descripción sobre los dos tipos principales de repositorios existentes: institucionales, temáticos, de datasets, de tesis doctorales, de revistas abiertas, así como bibliotecas digitales.

Los repositorios temáticos se definen por su especialización en documentos científicos de a un área del saber concreta. Siendo fieles a sus orígenes, los repositorios temáticos se alimentan esencialmente de *e-prints*, es decir, de versiones de artículos publicados en revistas, ya sean anteriores (preprints) o posteriores a su publicación (postprints).

Por su parte, los repositorios institucionales, más numerosos en conjunto, aunque de origen más reciente que los citados repositorios temáticos, surgen con el fin de recoger y conservar la producción científica de una institución. Los repositorios

institucionales suelen incluir una mayor variedad de tipologías documentales, y aportan varios beneficios esenciales: la posibilidad de *administrar* la documentación electrónica de la institución, antes dispersa en múltiples fuentes; una mayor *eficiencia* en esta administración, aunque su puesta en marcha suponga unos costes iniciales; un aumento de la *difusión* de los contenidos generados por los investigadores de la institución y, por lo tanto, de la *visibilidad* de la misma y de sus investigadores; y, por último una respuesta al sistema tradicional de publicación científica, que permite plantear una posible disminución de los costes de acceso a los contenidos científicos disponibles hasta entonces solo a través de las editoriales y distribuidoras comerciales de información científica.

El capítulo incluye una selección de ejemplos destacados de repositorios de todos los tipos indicados, y también recoge algunos de los principales proyectos e iniciativas, en distintos países o a nivel internacional, dirigidos a la creación y mantenimiento de proveedores de datos y de servicios o a la promoción de la Iniciativa de Archivos Abiertos.

1.4.4. Los proveedores de servicios

En cuanto a las iniciativas de proveedores de servicios, si bien menos numerosos, que los proveedores de datos, pueden distinguirse diferentes tipos en su enumeración. En primer lugar, se han destacado los proveedores de servicios generalistas (es decir, que recolectan proveedores de datos independientemente de las disciplinas científicas que abarquen) y de ámbito internacional, como OAlster, ScientificCommons, BASE o ARC, además del buscador de repositorios OpenArchives.eu, y de otros servicios que no solo recopilan contenidos OAI, sino también, otro tipo de proveedores de información científica en formato digital, como Scirus, Infomine o Google Scholar (también conocido como Google Académico).

Más centrados en áreas temáticas concretas, se han destacado proveedores de servicios que recopilan fuentes de áreas como la biblioteconomía y la documentación; las ciencias físicas o la ingeniería; y las ciencias naturales. Un tercer grupo de proveedores de servicios lo constituyen aquellos dedicados a recopilar registros de proveedores de datos de un determinado país o región geográfica. Asimismo, se describen proveedores de servicios especializados en proveedores de datos de recursos educativos y otros en el área de patrimonio cultural.

El capítulo 5 concluye con dos tipos de proveedores de servicios: por una parte; los implantados en una institución, para ofrecer un acceso integrado a los distintos repositorios o proveedores de datos de la propia institución; y por otra, los proveedores de servicios mantenidos por consorcios de bibliotecas.

1.4.5. La situación en Latinoamérica

El capítulo 6, escrito por Laureano Felipe Gómez-Dueñas, muestra el estado de la cuestión del movimiento de archivos abiertos en Latinoamérica. El capítulo comienza con un panorama genérico de la situación latinoamericana en cuanto a la adopción del

protocolo, las razones que lo han propiciado (especialmente, el aumento de la producción científica en los centros de investigación y las universidades, junto al incremento del interés hacia las publicaciones de otros investigadores latinoamericanos); así como en los problemas propios de la región en el desarrollo del proyectos OAI. Se incluye además un análisis detallado de los repositorios elaborados en países de la región, cuyo número total de repositorios asciende a 150.

A continuación, se abordan los diferentes proyectos sobre archivos abiertos desarrollados en diversos países latinoamericanos: Argentina, Brasil, Colombia, Chile, México y Venezuela. De todos ellos, son Brasil y Argentina los países con mayor número de iniciativas y de mayor envergadura, debido en buena parte al apoyo institucional recibido.

El capítulo contiene también una descripción de varios programas creados en la región para la implantación de proveedores de datos, y por último, se realiza una explicación sobre los principales proveedores de servicios de la región: entre los que se pueden encontrar repertorios de artículos de revistas, proveedores de tesis doctorales, buscadores de recursos internacionales de acceso abierto, proveedores de servicios específico por áreas científicas, además de dos proveedores de servicios centrados en recolectar registros de proveedores de datos nacionales.

1.4.6. La situación en España

El capítulo 7 está dedicado al desarrollo de OAI-PMH en España, el cual, según los datos de los registros más importantes, es uno de los diez países con mayor número de proveedores de datos, con cerca de treinta repositorios. Aunque ya en 2003 tuvieron origen algunas iniciativas pioneras, como TDX/TDR, ha sido en 2006 cuando han surgido buena parte de las iniciativas actualmente en funcionamiento.

La mayoría de las iniciativas españolas tienen su origen en las universidades, por lo que predominan los repositorios institucionales (al margen de la participación en repositorios temáticos como el citado E-LIS o Artxiker). Estos repositorios están en gran parte concentrados en las comunidades autónomas de Cataluña y Madrid, aunque no faltan desarrollos en otros puntos del país, como los de las universidades de A Coruña, Navarra, Alicante y Salamanca. Son especialmente importantes dos proyectos fruto de consorcios de universidades: el citado TDX/TDR, que recoge buena parte de las tesis defendidas en 17 universidades españolas (originalmente las universidades catalanas, a las que se han añadido otras); y el proyecto e-Ciencia, mantenido por el Consorcio Madroño de bibliotecas universitarias públicas en la Comunidad de Madrid, la Universidad Nacional de Educación a Distancia y el Consejo Superior de Investigaciones Científicas, que se materializa en un proveedor de servicios que actúa sobre los diferentes repositorios de acceso abierto desarrollados por cada una de las instituciones que participan en el proyecto.

En relación con las revistas electrónicas de acceso abierto publicadas en España se incluyen tanto portales o servicios de acceso a revistas como títulos individuales. También se incluyen diversos proveedores de datos financiados por instituciones públicas, como la Generalitat de Catalunya, el Ministerio de Cultura, o el Consejo Superior de Investigaciones Científicas.

Por último, el capítulo recoge los escasos proveedores de servicios en España, de ámbito nacional, regional o institucional, y financiados con fondos públicos o privados.

1.4.7. Herramientas de software para crear proveedores de datos y proveedores de servicios

Tanto para poner en marcha proveedores de datos como proveedores de servicios o agregadores, es importante conocer cuáles son las alternativas de software entre las que elegir. Aunque en los capítulos anteriores se dan algunas indicaciones, se ha considerado oportuno dedicar el capítulo 8, “Herramientas de software para OAI-PMH”, a elaborar una descripción de las principales aplicaciones empleadas en la actualidad.

Tras una primera introducción sobre la diferencia entre software libre, software de código abierto y software propietarios, se muestran las principales herramientas de software libre empleadas para la creación de repositorios, la generación de proveedores de datos que no son repositorios en sí, la recogida de metadatos mediante recolectores y proveedores de servicios; y, por último, otras herramientas que soportan OAI-PMH.

En primer lugar, se describen las principales herramientas de software para la creación de repositorios. Básicamente, todas ellas deben cumplir algunos requisitos:

- soportar el flujo documental: introducción de documentos; rechazo o aceptación de los mismos por los administradores del repositorio; edición de los -metadatos sobre los recursos, etc.
- ofrecer para estas acciones diferentes niveles de acceso: usuario, autor y administrador, así como permitir la creación de grupos de usuarios que tengan el mismo nivel de acceso.
- crear colecciones de materiales, que puedan ser archivados en más de un formato.
- permitir la navegación, búsqueda, recuperación, visualización y descarga de esos materiales.

Este bloque se centra en una serie de aplicaciones que cumplen varios requisitos: soportan la última versión de OAI-PMH, la 2.0; son software de código abierto, es decir, este código es libremente modificable, y pueden ser descargadas pública y gratuitamente; permiten crear repositorios genéricos, de cualquier tipo; y son los más utilizados en su conjunto en el mundo. Estos son, por orden alfabético: Archimede, ARNO, CDS Invenio, DSpace, Fedora, GNU EPrints, MyCoRe y OPUS. A ellos se han añadido otras herramientas para usos más específicos, como ETD-db y DoKs (para tesis) o Kepler (para la creación de archivos personales).

Para cada uno de ellos, se han recogido algunos datos básicos: la institución que mantiene cada repositorio, el tipo de documentos recogidos, el software necesario para su implantación, las opciones para la introducción de los registros y para su administración técnica y documental, las opciones de búsqueda y recuperación, los esquemas de metadatos empleados para la descripción documental y, si las tiene, las opciones de exportación de datos.

El segundo bloque recoge otras herramientas capaces de generar proveedores de datos a partir de distintas estructuras de almacenamiento de información, como bases de datos y sistemas de ficheros XML, así como de pasarelas Z39.50.

En el bloque de herramientas para la creación de proveedores de servicios, se recogen algunas iniciativas de especial interés, como Arc, de la Universidad de Old Dominion, en Estados Unidos, o PKP Open Archives Harvester, una herramienta generada por el proyecto Public Knowledge Project (PKP) de las universidades canadienses Simon Fraser y Columbia. También se incluye la descripción de otras herramientas menos extendidas, como: Perl Harvester, Java OAI Harvester o VB OAI Harvest.

El capítulo concluye con la mención de una serie de aplicaciones que también han adoptado el protocolo OAI-PMH con distintas finalidades: la monitorización de repositorios que cumplan el protocolo, extensiones para los sistemas de gestión de contenidos; herramientas para generar publicaciones en acceso abierto como revistas y actas de conferencias; y otras herramientas para permitir el acceso a repositorios estáticos o para convertir metadatos en otros formatos al formato mínimo especificado por el protocolo, Dublin Core sin cualificar.

1.5. ALGUNAS FORMAS DE ACCEDER A LA OBRA

La organización de este trabajo está pensada para diferentes públicos. El lector que está interesado sólo en tener una visión introductoria del movimiento de acceso abierto y la Iniciativa de Archivos Abiertos, podrá adquirirla mediante la lectura de los dos primeros capítulos, y si desea obtener una visión general de la situación de dichos movimientos en todo el mundo, y específicamente en España y Latinoamérica, se recomienda la lectura de los capítulos 5, 6 y 7. Un lector que ya tenga un cierto conocimiento de las ideas, fundamentos, procesos del acceso abierto y los archivos abiertos y cómo éstos se ven afectados por el uso del protocolo para la recolección de metadatos (OAI-PMH), pero quiera profundizar en aspectos más técnicos del mismo, podrá obtenerlo mediante la lectura de los capítulos tercero y cuarto, y si además desea conocer las herramientas software que cumplen con el protocolo, podrá leer el capítulo 8. Aunque estos capítulos son más técnicos (especialmente el 3 y el 4, basados en las especificaciones y guías de implementación de la OAI), hasta donde ha sido posible se ha tratado de hacerlos legibles y de poner explicaciones y ejemplos simplificados siempre intentando no pasar el límite en el que se pierda el rigor técnico.

Este campo tiene, además, la particularidad de su juventud y cambio continuo. Como continuamente surgen nuevos proyectos y desarrollos, y se actualizan los ya en funcionamiento, la bibliografía existente queda rápidamente superada por la realidad. Esto hace que gran parte de las referencias que aparecen en este trabajo sean páginas web de las instituciones implicadas, publicaciones electrónicas o documentos que están disponibles en Internet.

En todos los casos, dada la gran variedad de términos empleados, inevitable en un tema como el tratado, en continuo desarrollo y evolución, se ha considerado útil añadir

algunas secciones especiales, como una recopilación de recursos sobre el movimiento *open access* y la Iniciativa de Archivos Abiertos, que constituye el capítulo 9, así como un glosario y una lista de siglas, así como un índice de las figuras incluidas.