

E. RECURSOS DE INFORMACIÓN

Informe de situación

Nuevas fuentes de información académica

Por **Lluís Codina**

Codina, Lluís. "Nuevas fuentes de información académica".
Anuario ThinkEPI, 2010, v. 4, pp. 139-142



Resumen: Se describen las características básicas de uso y consumo de la información científica, técnica y médica (STM), y se presentan 3 recursos de información creados recientemente: DeepDyve; Science Research, de Deep Web Technologies; y WorldCat, de OCLC. También se mencionan Q-Sensei, Base, Scitopia y Lumifi.

Palabras clave: Fuentes de información, Recursos, Información científica, Información académica, Información técnica, STM.

Title: *New sources of scholarly information*

Abstract: The basic characteristics of use and consumption of scientific, technical and medical (STM) information are described, and three newly created information resources are presented: DeepDyve; Science Research from Deep Web Technologies; and WorldCat, from OCLC. Also mentioned are the services Q-Sensei, Base,

Scitopia and Lumifi.

Keywords: Information sources, Resources, Scientific information, Scholarly information, Technical information, STM.

1. La comunidad STM

Un rasgo común a ciertas comunidades profesionales es que para desarrollar su trabajo necesitan obtener información científica, técnica o médica (STM a partir de ahora, por seguir las conocidas siglas inglesas y no inventar otras nuevas). Forman un grupo característico, dentro del grupo más amplio de profesiones intensivas en información, que suele llevar a cabo su trabajo basándose principalmente en la denominada información cognitiva (por contraste con la información factual, administrativa, etc.).

Ahora bien, a efectos de este trabajo caracterizaremos también a estos profesionales del sector STM bajo el marco de dos *presiones* y de tres *necesidades*.

En cuanto a las presiones, la primera consiste en que los miembros de tales comunidades deben utilizar como uno de los inputs de su trabajo *información de calidad*.

Esto significa que únicamente pueden llevar a cabo su labor manejando informaciones que estén de alguna forma certificadas o garantizadas mediante esquemas de evaluación comúnmen-

te aceptados (en general esquemas tipo *peer review*).

La segunda presión es una consecuencia lógica de la primera: se espera de estos profesionales que sean capaces de producir *nuevas informaciones* o nuevos conocimientos (a su vez susceptibles de ser evaluados).

Las presiones anteriores generan necesidades paralelas. La primera consiste en disponer de *yacimientos de información* donde poder acudir cuando necesitan nuevos inputs de información de calidad para su trabajo (p. e. para una nueva investigación). Este acopio de información debe hacerse de forma continuada, y de poco servirá si no se dispone de medios para gestionarla de forma muy eficiente cada vez que debe ser (re)utilizada.

Pero la memoria humana necesita recurrir de forma sistemática a instrumentos auxiliares que se utilizan como memorias exosomáticas. O sea, esas informaciones, a veces trabajosamente obtenidas, no serán memorizadas de forma indeleble sin recurrir a herramientas de almacenamiento externas.

Aunque tendemos a retener conceptos, habi-

lidades, teorías e ideas sin que se sepa de ningún límite claro para ello, en cambio no (necesariamente) tendemos a recordarlo todo a nivel de detalle, ni a recordarlo justo en el momento en el que lo necesitamos; tampoco acertamos siempre, ni mucho menos, en detectar las relaciones entre las informaciones recopiladas y los valores que asumen las variables con las que necesitamos trabajar.

Elementos de la información STM

P1 Consumidores
P2 Productores
N1 Yacimientos
N2 Sistemas de gestión
N3 Medios de divulgación

En definitiva: identificar en un momento determinado (t_1) una información o un conjunto de datos obtenidos a veces después de un trabajo intenso de búsqueda en los yacimientos a los que nos hemos referido, no garantiza que, en un momento posterior (t_2) acertemos a reutilizarlo.

Todo esto nos lleva a la segunda necesidad: *sistemas para gestionar* aquella información valiosa que eventualmente deberá actuar como input en los diferentes proyectos. A diferencia de los yacimientos de información, que son de ámbito social, estos sistemas han de ser ante todo capaces de adaptarse a perfiles personales y eventualmente de grupo.

Por último, los profesionales de la STM necesitan divulgar su producción. De nada le sirve a un joven doctor, o a un profesor de universidad haber producido un gran trabajo académico, si no dispone de herramientas para divulgarlo de forma eficiente en la era de la Web. De hecho, es una obligación o una responsabilidad legal para ciertos miembros de esta comunidad, como los profesores de universidad, cuya carrera académica está ligada a su éxito para divulgar su producción científica.

Aunque para muchos todavía la Web es sinónimo de desorden y de información de baja fiabilidad, lo cierto es que en los últimos años se ha convertido en la principal fuente de todo lo contrario, es decir en el lugar privilegiado donde los profesionales del sector STM puede encontrar recursos para:

a) Obtener información con el objetivo de producir nuevas informaciones y nuevo conocimiento.

b) Gestionar la información obtenida para poder reutilizarla.

c) Difundir el nuevo conocimiento de modo que pueda ser utilizado como input de información por parte de otros investigadores.

2. Tres recursos STM destacados

En lo que sigue se presentan algunas características de tres de los nuevos buscadores que, en nuestra opinión, merecen ser tenidos en cuenta por el conjunto global de sus prestaciones y por su capacidad para convertirse a medio plazo en buscadores de referencia, algunos de ellos con ventaja al menos en algunos aspectos sobre los dos grandes "clásicos" (*Scholar* y *Scirus*). La elección es inevitablemente subjetiva, pero esperamos que también sea razonable.

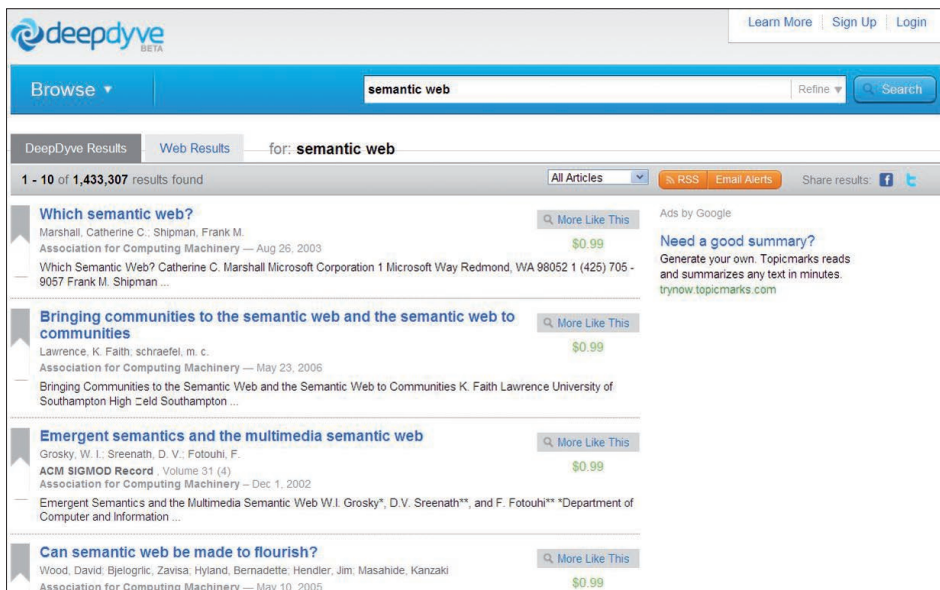
2.1. DeepDyve

<http://www.deepdyve.com>

DeepDyve es una empresa norteamericana fundada en 2005 por científicos procedentes del proyecto del genoma humano.

Por lo que respecta a su contenido (artículos de revistas académicas de prestigio) es muy parecido a las bases de datos académicas convencionales, como por ejemplo, *Science Direct*, pero con una importante diferencia: está explorando un nuevo modelo de negocio, o si se prefiere un acercamiento diferente al tradicional. Permite la búsqueda en miles de títulos de las revistas más importantes de diversas áreas, incluyendo las ciencias, las humanidades y las ciencias sociales. En total, contiene unos 32 millones de artículos. Ahora bien, en lugar de tener un enfoque de suscripción orientado a grandes corporaciones, se dirige al usuario final incluyendo una modalidad de suscripción gratuita con un precio unitario máximo por artículo completo de 1 US\$. Esto es entre 20 y 30 veces menos de lo que suelen pedir las bases de datos académicas tradicionales en la modalidad equivalente (pago por artículo). A cambio, no se puede comprar el artículo, sino que únicamente se "alquila" (sólo puede leerse en pantalla, no puede descargarse). En este sentido, sería un acercamiento al modelo de los *e-books*. El alquiler puede ser de 24 horas (suscripción gratuita), siete días o permanente (estos dos últimos en las dos modalidades de pago).

Es importante porque ofrece un nuevo modelo de negocio que puede impactar favorablemente en entornos donde no existe la posibilidad de afrontar las suscripciones de las tradicionales bases de datos, o cuando, una corporación (empresa, hospital, biblioteca, etc.) o un investigador individual necesite puntualmente artículos



DeepDyve. Vista parcial de una página de resultados

de revistas no suscritas. Cabe destacar que, incluso la modalidad de suscripción más cara es parecida o inferior al precio típico de la suscripción a una sola revista académica, con la diferencia que se puede acceder al contenido de miles de ellas sin ningún límite ni de cantidad de artículos accesibles ni de tiempo de alquiler.

Incluye artículos gratuitos procedentes de revistas open access.

2.2. Science Research

<http://scienceresearch.com>

Deep Web Technologies es una empresa norteamericana fundada en 2002, especializada en búsquedas y gestión del conocimiento. Ha producido diversos buscadores especializados en ciencia. El que comentamos aquí es el más genérico de ellos.

Es un sistema de búsqueda federada, es decir, que en lugar de indizar la información y tenerla en su base de datos, como hace Scirus, por ejemplo, lo que hace es enviar la misma pregunta a centenares de colecciones académicas cada vez que un usuario pregunta. Las colecciones pueden ser bases

de datos, repositorios u otros buscadores. En estos momentos, Science Research es capaz de consultar unas 400 de tales colecciones, por lo cual el volumen de resultados que puede ofrecer para casi cualquier búsqueda es enorme, aunque también es muy grande el número de duplicados que aparecen a veces en sus páginas de resultados. A pesar de que el sistema intenta eliminarlos, no en todas las búsquedas tiene el mismo éxito.

Es importante porque ofrece la búsqueda

federada más ambiciosa, con el objetivo de ser de hecho un buscador universal. De refinar su tecnología, debería aumentar la exhaustividad del buscador.

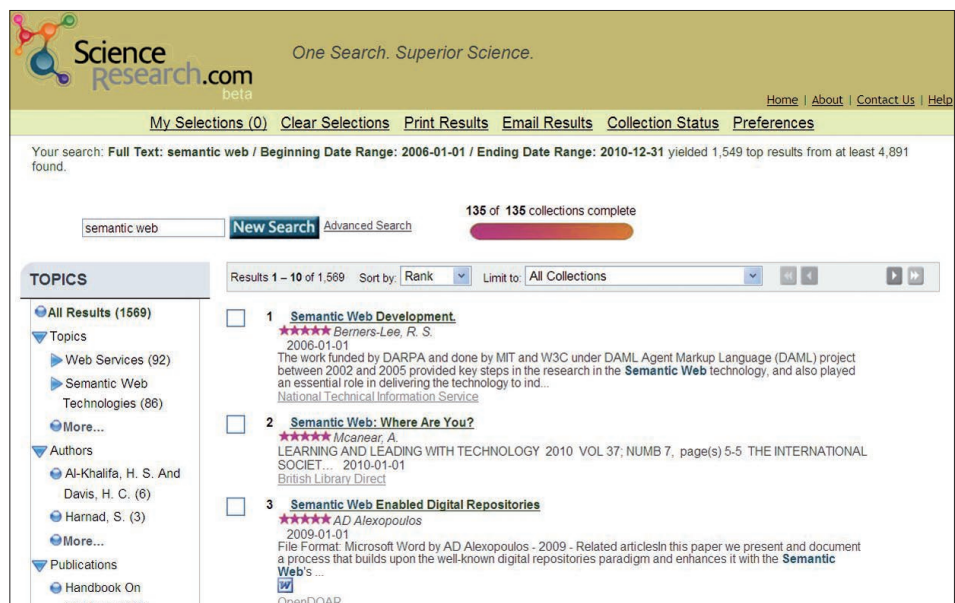
2.3. WorldCat

<http://www.worldcat.org>

OCLC es la organización más importante en servicios para bibliotecas universitarias.

Afirma ser el catálogo más grande de materiales académicos con más 1.500 millones de documentos.

Es un buscador tipo recolector, es decir, intro-



Science Research. Vista parcial de una página de resultados



WorldCat. Vista parcial de una página de resultados

duce en su índice registros producidos por diferentes productores de bases de datos, entre ellos los procedentes de catálogos producidos por centenares de bibliotecas de todo el mundo, así como registros de repositorios digitales también de origen universitario. *WorldCat* no es nuevo, pero sí lo es la integración de buscadores como *OAister* y sus nuevas prestaciones tipo web 2.0.

3. Otros recursos académicos

En cuanto a abundancia de recursos, la información de tipo académico está pasando por una excelente etapa. A continuación se recogen otros buscadores y sistemas de información especializados, de creación más o menos reciente. Qué duda cabe que la selección natural actuará sobre ellos, de modo que algunos dejarán de existir en el medio plazo, pero sin duda algunos otros han llegado para quedarse, de modo que los miembros de la comunidad STM tienen por ahora un panorama del que podemos decir que, por lo menos, es interesante.

Recursos recientes de información académica

Q-Sensei

<http://www.qsensei.com>

Buscador de la empresa del mismo nombre. Resultado de la fusión de una compañía alemana y norteamericana. Afirma indizar 24 millones de artículos y 11,5 millones de libros. Combina la indización y la búsqueda federada.

Base

<http://www.base-search.net>

Buscador tipo recolector desarrollado por la biblioteca de la *Universidad de Bielefeld* (Alemania). Proporciona acceso a casi 24 millones de documentos procedentes de casi 1.600 fuentes. Utiliza tecnología de la empresa *FAST* (adquirida por *Microsoft*), la misma que utiliza *Scirus*.

Scitopia

<http://www.scitopia.org>

Sistema de búsqueda federada de la misma empresa creadora de *Science Research*. Está especializado en física, tecnología y patentes.

Lumifi

<http://www.lumifi.com>

Es una combinación de buscador, base de datos personal, sistema de análisis de la información y sistema colaborativo. Es el modelo de lo que podrían ser los auténticos sistemas de información académicos de la Web 2.0 (o de la 3.0). Su único problema es que para las búsquedas utiliza *Gigablast*, un motor sumamente modesto.

4. Conclusiones

Como es sabido, además de los sistemas que hemos presentado aquí y de *Google Scholar* y *Scirus*, disponemos de numerosos portales y bases de datos de editoriales académicas como los de *Elsevier* (*Science Direct*, *Scopus*) o *Springer*; de bases de datos de productores como *Thomson (ISI)*, de bases de datos de distribuidores de bases de datos como *Ebsco*, etc.

Es decir, nosotros nos hemos limitado a presentar aquí, en función de su novedad, una pequeña muestra de la oferta en recursos de información para la comunidad STM. Lo cierto es que acceder a grandes cantidades de información, y hacerlo a la vez de una forma selectiva, ya no es un gran problema.

Otra cosa es seleccionar en cada caso el mejor (o mejores sistemas), pero sobre todo, el problema es qué hacer con lo que encuentra el usuario de estos sistemas. En teoría hay también numerosas soluciones, desde sistemas comerciales como *RefWorks* hasta sistemas abiertos como *Zotero* o *Mendeley*, pero aún no parece estar bien resuelto ni mucho menos este lado del problema.