

Red ARACNE: retos y objetivos de un proyecto de coordinación en letras hispánicas digitales

CONSOLACIÓN BARANDA
Instituto Universitario Menéndez Pidal (Universidad Complutense)
cbaranda@ucm.es

EDUARDO RODRÍGUEZ
Universidade da Coruña
eduardo.rodriguez.lopez@udc.es

Los especialistas en humanidades hemos acogido con gran entusiasmo las posibilidades que ofrecen las herramientas digitales para la investigación en nuestras áreas de trabajo. Como consecuencia de ello se ha producido tal multiplicación de páginas web relacionadas con la filología que nos puede conducir a un efecto perverso: la pérdida de cada uno de los proyectos en el océano de la red, la inanidad de cada esfuerzo particular.

De la conciencia de este problema surgió la organización de un seminario en junio del año pasado: el *I Seminario Internacional sobre Bibliotecas digitales y Bases de datos especializadas para la investigación en Literaturas Hispánicas (BIDESLITE)* que se realizó en Madrid los días 4 y 5 de julio de 2011. Lo dirigió Mercedes Fernández Valladares y lo organizó el *Grupo de Estudios de Prosa hispánica bajomedieval y renacentista* de la Universidad Complutense de Madrid con financiación del MICINN por medio de una Acción Complementaria (FFI2011-12620-E) y la ayuda del *Instituto Universitario Menéndez Pidal*. Las actas del *Seminario* están publicadas y disponibles en acceso abierto a través de *E-Prints*, el repositorio institucional de la Universidad Complutense <<http://eprints.ucm.es/21207/>>.

Las conclusiones de este seminario se podrían resumir en tres puntos:

- La necesidad de potenciar las estrategias de visibilidad y difusión de los recursos generados al amparo de este tipo de trabajos.
- Convertir esa visibilidad en elemento de identificación y desarrollo de líneas de investigación con base tecnológica de calidad, fortaleciendo la coordinación y su incidencia en el sistema público de investigación.
- La conveniencia de generar instrumentos de enlace y mejora de la recuperación de datos a partir de las bases y bibliotecas digitales existentes. El desarrollo de recursos comunes y herramientas de diseño específico con objeto de proporcionar información de calidad para crear otras herramientas en el área.

La confluencia de estas ideas dio lugar a la percepción, compartida unánimemente entre los ocho grupos participantes, de estar trabajando en proyectos que, aun siendo diferentes en contenidos, objetivos, metodología, géneros literarios o períodos temporales, habían tenido que solventar problemas muy parecidos ofreciendo soluciones –muy dispares en algunos casos– de las que convendría extraer un rédito común.

Como resultado del intercambio de reflexiones se presentó una solicitud de proyecto del conjunto de grupos coordinados en forma de Ayuda Complementaria a 30 meses, con el Dr. Pedro Ruiz como I.P. Estas son las bases de la Red Aracne que presentamos ahora, en la que se integran 6 de los ocho grupos que participaron en el Seminario. Por razones técnicas, los equipos de *CANCIONERO VIRTUAL* de la Universidad de Liverpool y *NICANTO* de la Universidad de Burdeos III no pudieron incorporarse a esta solicitud¹.

La red Aracne está integrada por los siguientes proyectos de investigación:

BIESES. BIBLIOTECA DE ESCRITORAS ESPAÑOLAS (<http://www.bieses.net/>)

En la base de datos de BIESES se recogen todas las referencias primarias y secundarias de las escritoras españolas (en sentido amplio) hasta 1800; el arco temporal enlaza así con la bio-bibliografía de escritoras del siglo XIX de Carmen Simón Palmer.

La información en BIESES está construida y seleccionada en función de una metodología de estudio específica y, aunque útil en otros aspectos y para otros intereses, en un futuro deberá ser capaz de ofrecer los datos necesarios para responder a las preguntas que se formulan desde esa opción crítica. Algunas de estas preguntas implícitas pueden ser las siguientes. En cuanto a la identidad biográfica o social de las autoras: cuántas escritoras hay en el período, cómo se distribuyen cronológica y geográficamente, qué coincidencias vitales hay entre ellas que puedan sugerir posibles conexiones o relaciones; cuántas de estas escritoras pertenecieron a órdenes religiosas, qué órdenes, en qué conventos, cuándo. En lo que se refiere a sus obras: cuántas y de qué tipo están en manuscritos, qué se edita, dónde, cuándo, cómo, por medio de qué agentes culturales o mediadores. Por último, BIESES aún puede responder a otras cuestiones de recepción, a través del campo de las menciones coetáneas que recogen citas de otros autores, de las cartas intercambiadas, de los paratextos o del intercambio poético; también se revisan inventarios de bibliotecas con el fin de averiguar qué autoras son leídas y por quién o por qué autoras que se espera sean citadas no lo son en algunos contextos.

BIDISO. BIBLIOTECA DIGITAL SIGLO DE ORO (<http://www.bidiso.es/sielae/>)

El grupo SIELAE (Seminario Interdisciplinar para el Estudio de la Literatura Áurea Española) ha propiciado desde 1992 la creación sucesiva de varias bases de datos y bibliotecas digitales en torno a cuatro proyectos:

¹ Aunque por razones históricas *NICANTO* no ha podido realizar todavía el desarrollo informático, está proporcionando material inédito y desconocido o poco conocido a algunos grupos de la red-Aracne; es ejemplo de trabajo útil y fundamental, a pesar de no contar con herramientas informáticas actuales.

- Emblemática.
- Relaciones de sucesos.
- Polianteas y libros de erudición.
- Inventarios y Bibliotecas del Siglo de Oro (IBSO).

El proyecto se presenta en este *I Congreso Internacional de la HDH*, por lo que remitimos a su intervención.

BSF. BIBLIOTECA SAAVEDRA FAJARDO DE PENSAMIENTO POLÍTICO HISPÁNICO (<http://saavedrafajardo.um.es>)

La labor fundamental de la Biblioteca Saavedra Fajardo es centralizar una serie de herramientas útiles para el investigador dedicado al pensamiento político hispano, así como presentar un catálogo extenso y materiales adicionales de forma ordenada, para acceder con facilidad a recursos propios y externos. Es, en este momento, la que cuenta con mayor volumen de información especializada de nuestra asociación.

Atiende al estudio de las fuentes hispánicas y a la producción filosófica en nuestro idioma y está integrada por cuatro colecciones, una base de datos bio-bibliográfica de fuentes primarias y secundarias sobre pensadores hispánicos, una hemeroteca digital con 11 títulos de revistas especializadas y una sección de Tribuna con artículos breves y pequeños ensayos, abierta a la colaboración externa. Ofrece documentación biográfica y bibliográfica, edita fuentes de la historia del pensamiento hispánico, en versión de fichas, con índices descriptivos de cada página, y en versión pdf, con textos completos, y pone en la red artículos y trabajos de investigación sobre filosofía y pensamiento español de todas las épocas.

El Centro de Documentación Gonzalo Díaz y María Dolores Abad vertebró el funcionamiento de la BSF; es un ingente repertorio biobibliográfico de pensadores de la filosofía española que tiene como núcleo la magna obra de Gonzalo Díaz *Hombres y documentos de la filosofía española*.

A día de hoy hay publicados en el GDDA los repertorios de 2.666 pensadores hispanos, con más de 52.000 obras y más de 33.000 estudios referenciados. Cuando una obra o estudio está disponible en Internet, basta con seleccionar la referencia en el Centro de Documentación para acceder a la misma.

CLARISEL. BASES DE DATOS (<http://clarisel.unizar.es/>)

Es un portal que reúne tres bases independientes.

Los trabajos incluidos en *AMADÍS* pretenden arrojar nuevas luces sobre la producción literaria caballeresca compuesta en diferentes modalidades genéricas desde la Edad Media hasta el siglo XVII, aparte de incluir estudios posteriores que tengan como referencia intertextual obras anteriores.

Además, si bien el ámbito de trabajo es fundamentalmente filológico, se incorporan entradas interdisciplinarias –de historia, derecho, arte, folclore, antropología, etc.–, que puedan ayudar a interpretar y contextualizar la materia objeto de atención en un ámbito ibérico. Esto implica que incluye referencias temáticas, como los trabajos sobre los “rieptos” o sobre tácticas militares, órde-

nes militares, etc., referidas a textos castellanos, catalanes, gallegoportugueses y portugueses. El término de partida es 1998 por ser la fecha de los inicios oficiales de la página, teniendo en cuenta, además, que en el 2000 se publicó la bibliografía sobre los libros de caballerías de Daniel Eisenberg y M^a Carmen Marín.

SENDEBAR recoge y sintetiza la bibliografía sobre el cuento medieval hispánico a partir de 2005. Incluye ediciones y estudios sobre cuentos escritos en castellano, catalán y gallego-portugués, así como en otras lenguas occidentales y orientales (latín, árabe, hebreo, portugués, francés, italiano, etc.), siempre que se trate de obras que mantengan conexiones, directas o indirectas, con producciones hispanas. Con idénticas restricciones, se incorporan publicaciones sobre el relato folclórico y sus métodos de análisis, dada la interrelación entre el cuento escrito en la Edad Media y la tradición oral.

HEREDIA está dedicada a la literatura escrita en Aragón durante la Edad Media y el siglo XVI, con independencia también de la lengua; así, en ella se incorpora información sobre Pedro Alfonso, literatura aljamiado-morisca, escritos en aragonés o en catalán, etc. Como la anterior, fue creada en el año 2005, si bien surgió ya con la vocación de incluir, poco a poco, referencias impresas anteriores; cuenta en la actualidad con más de 1.230 entradas.

En todos los casos, además de bibliografía y resúmenes de contenido, se ofrecen los enlaces electrónicos que están disponibles tanto a los textos como a las revistas mencionadas.

DIALOGYCA BDDH. BIBLIOTECA DIGITAL DE DIÁLOGO HISPÁNICO (<http://www.ucm.es/info/dialogycabddh/>)

Se propone recopilar, analizar y poner a disposición de los investigadores el *corpus* de diálogos hispánicos, desde sus orígenes hasta la actualidad, mediante una base de datos y una biblioteca digital. De acuerdo con estos planteamientos, hemos elaborado una base de datos que proporciona información sobre cada uno de los diálogos, ahora preferentemente de los siglos XV a XVII, pues en esta primera fase nos hemos concentrado en la etapa dorada del género, aunque eso no es obstáculo para introducir fichas analíticas de otros períodos.

Atendemos a los siguientes aspectos: el autor y otros responsables secundarios de la obra, las características literarias y modalidades de difusión, los personajes, las materias principales, los temas y formas de enunciación y lengua, la bibliografía primaria –con la localización y referencia de los testimonios manuscritos e impresos– y la bibliografía secundaria específica sobre cada diálogo.

Pero además, *Dialogyca BDDH* ofrece la reproducción completa del testimonio más relevante para su transmisión –de ahí el carácter de Biblioteca o colección digital–, en unos casos –hasta ahora la mayoría– digitalizado por nosotros y presentado en un documento pdf con permiso de difusión de la biblioteca que lo custodia; en otros mediante enlaces a las digitalizaciones realizadas por otras instituciones, si las hay.

De este modo, cada ficha encierra de forma organizada un estado de la cuestión sobre un diálogo concreto junto con su texto, y la Biblioteca digital proporcionará una visión de conjunto sobre el género: permitirá conocer mejor el número de autores y obras, ayudará a percibir la desigual atención crítica recibida por cada uno de ellos y favorecerá la posibilidad de establecer redes

de relaciones sincrónicas y diacrónicas entre los textos, sus temas, sus lugares o fechas de impresión, sus personajes, sus fuentes, etcétera.

Este diseño, basado en el concepto de *obra*, marca una diferencia sustancial con otras bases de datos de tipo catalográfico o bibliográfico y de carácter generalista, cuya unidad articuladora suele ser el documento –manuscrito o impreso– y atienden preferentemente al lugar o institución en que se ha conservado; por el contrario, el concepto de *obra* como unidad articuladora nos permite agrupar testimonios y documentos de diversa tipología, conservados en bibliotecas muy dispersas y recoger de ese modo, junto a la información literaria la tradición textual de cada diálogo.

PHEBO. POESÍA HISPÁNICA EN EL BAJO BARROCO (<http://phebo.es>)

El proyecto *PHEBO* (Poesía Hispánica en el Bajo Barroco) surgió a mediados del 2010 en el seno del equipo que desarrollaba el estudio “La poesía de bajo barroco: repertorio y categorías”. El objetivo que se plantea es sentar las bases para una cartografía de la poesía de este período, tan poco atendido por la crítica y la historia literarias, a partir de dos elementos básicos: un repertorio de los impresos con materia poética entre 1650 y 1750, de un lado, y una revisión de los conceptos y categorías aplicados al análisis de la poesía áurea, de otro. La base de datos persigue el establecimiento de las herramientas básicas para el análisis, estudio y valoración de la práctica poética en el Bajo Barroco mediante la elaboración del repertorio de impresos poéticos entre 1650 y 1750, la edición de los textos y la recopilación de la bibliografía secundaria especializada. El elemento distintivo lo constituye la edición, esto es, la práctica fundamental de la filología y la base de todo estudio que pueda ser compartido.

Se decía al comienzo que somos conscientes de las diferencias en contenidos, objetivos, metodología, géneros literarios o períodos temporales de nuestros proyectos, pero también lo somos de que, en parte por ello, nos complementamos y también de que compartimos puntos de partida importantes. Como objeto de estudio, predominan los proyectos que ambicionan ofrecer un panorama de conjunto sobre un género literario: Cuento medieval, Libros de caballerías, Diálogo Literario, Emblemática, Relaciones de sucesos, Polianteas o libros de erudición y Poesía del bajo barroco.

Otros integran más de un género, pues han partido de diferentes criterios y tienen un carácter transversal: *BIESES*, *Heredia*, *IBSO* y *Saavedra Fajardo*. El abanico temporal se extiende desde la edad media hasta hoy, con especial presencia de los siglos XV, XVI y XVII en este momento.

Cada uno de los proyectos tiene objetivos claramente delimitados, ofrece información discriminada y organizada de acuerdo con esos objetivos, aporta valor añadido a sus respectivos ámbitos de estudio, recoge información altamente especializada y busca la utilidad pública, así como servir de referencia informativa en su campo. Todos integran el contexto de un modo u otro, recogen fuentes primarias y, en muchos casos, también secundarias; aunque trabajan fundamentalmente sobre impresos, también se encuentran presentes los manuscritos. Comparten también un sentido abarcador de las Humanidades Hispánicas, pues incluyen textos en las distintas lenguas peninsulares y también hispano-latinos, aljamiados, portugueses, etcétera.

Se trata de proyectos en marcha, con procesos de actualización permanentes, que han previsto la colaboración de investigadores en formación, junto con especialistas y colaboradores externos, de ahí que consideren importante hacer constar la autoría de cada una de las aportaciones, elemento fundamental como respaldo del rigor científico. Por supuesto, cuentan con financiación pública y con un soporte institucional que auspicia su estabilidad.

La Red Aracne se apoya en dos pilares: el carácter complementario de los proyectos que la integran y la concurrencia de criterios en el planteamiento científico de todos ellos. El objetivo último no es sumar la información, sino multiplicarla, porque se trata de poner en relación datos previamente procesados, elaborados y estructurados con rigor². Pretendemos, por una parte, rentabilizar los medios y el trabajo invertidos potenciando su utilización y conseguir visibilidad, pero también construir una vía de futuro inédita para las Humanidades Digitales Hispánicas. La red Aracne podrá servir como banco de pruebas y futuro punto de convergencia de otros grupos que trabajan en el campo de las Humanidades Digitales.

Para la creación de la red Aracne hemos elegido avanzar con dos núcleos de trabajo fundamentales:

1. Creación de un portal web que centralice y ordene la información sobre estos proyectos e instrumentos en el campo de las humanidades digitales aplicadas a las letras hispánicas, con objeto de que sirva de referencia, orientación, convergencia y principio de coordinación entre iniciativas, las ya existentes y las emergentes.
2. Paralelamente, la elaboración de un *metabusca* que ofrezca la posibilidad de una búsqueda simultánea en todas las plataformas participantes y ofrezca resultados coherentes y exhaustivos, aunque no necesariamente completos, de todas las informaciones que aparezcan en las distintas bases de datos. Este tipo de experiencia y las problemáticas, tanto técnicas como metodológicas, que se lleguen a identificar y resolver, servirán para crear pautas para la incorporación de nuevos proyectos a la red y también como ejemplo para los proyectos que quieran compartir los resultados, pero no la aplicación informática.

Además, de manera paralela o sucesiva, según los casos, nos proponemos avanzar en otros objetivos:

- Diseño de un modelo de criterios de evaluación de calidad técnica y científica de los productos resultantes, con la finalidad de homogeneizar protocolos y propiciar su aceptación para el reconocimiento y evaluación de los trabajos realizados.
- Desarrollo de una herramienta flexible de carácter interactivo orientada a la enseñanza y formación de los alumnos en los inicios de la carrera investigadora.

² A modo de prueba, Mercedes Fernández Valladares hizo una simulación buscando en *Dialogyca BDDH* datos relativos a Aragón que pudiesen complementar la base *Heredia* y la variedad de resultados fue superior a la que se podía prever: aparecieron autores que nacieron, murieron o desempeñaron actividades profesionales en Aragón, libros allí impresos o conservados en sus bibliotecas, e incluso personajes, interlocutores de algún diálogo, estaban relacionados con Aragón.

- Desarrollo de la ampliación del proyecto y búsqueda de convocatorias oficiales y fuentes de financiación (en especial en el marco europeo), para llevarlo a cabo.

Con ellos se busca establecer una sinergia que permita la confluencia de los equipos que inicialmente respaldan el proyecto más aquellos que en el futuro deseen sumarse a partir de la colaboración en actuaciones concretas y delimitadas que permitan una homologación de estrategias de investigación, una integración de resultados y una tecnología común.

En lo referente a los desarrollos tecnológicos, cuya responsabilidad ha sido asignada al Laboratorio de Bases de Datos de la Universidad de A Coruña (<http://lbd.udc.es>), en la actualidad se están avanzando en las tareas conducentes al cumplimiento de los dos objetivos fundamentales de proyecto, que son, la elaboración de un portal web para la difusión de la *Red* y su actividad, y la implementación de una herramienta de búsqueda que permita buscar en todas las bibliotecas digitales y bases de datos integradas en la misma.

Metabuscador de recursos digitales

El principal problema con el que nos encontramos en el desarrollo del *metabuscador* es la heterogeneidad de los repositorios digitales a integrar. Cada biblioteca digital o base de datos que forman parte la *Red* ha sido diseñada atendiendo a unos criterios específicos, propios del dominio de los recursos con los que trabaja, e implementada utilizando diferentes herramientas software que, además, muchas veces condicionan notablemente la propia estructuración de los datos catalogados. Además, la integración ha de hacerse de forma que no se modifiquen las bibliotecas digitales y bases de datos originales.

Este problema de la heterogeneidad no es específico del proyecto *Red ARACNE*, sino que afecta en mayor o menor medida a todos los proyectos de integración de colecciones documentales digitales. Existen diversas aproximaciones para construir sistemas que permitan integrar distintas fuentes de datos de cara a su interoperabilidad. La solución que mejor se adapta a las necesidades concretas de la *Red ARACNE* es la de utilizar el protocolo OAI-PMH (*Open Archives Initiative-Protocol Metadata Harvesting*, <http://www.openarchives.org/pmh/>) para la compartición de los metadatos descriptivos de los recursos digitales, sobre los que posteriormente se resolverán las búsquedas. Al tratarse de un estándar abierto y ampliamente utilizado en la actualidad por distintas iniciativas institucionales, la utilización del protocolo OAI-PMH permitirá también la difusión de los recursos digitales de la *Red ARACNE* a través de grandes repositorios de recursos digitales como Europeana (<http://www.europeana.eu>), promovido por la Unión Europea, o Hispana (<http://hispana.mcu.es>), del Ministerio de Educación, Cultura y Deporte.

Los recursos digitales se describirán utilizando el conjunto de 15 elementos básicos Dublin Core (*Dublin Core Metadata Initiative*, <http://dublincore.org/>), entre los que se encuentran el título, la autoría, una descripción, la temática, el idioma o el formato. Todos estos elementos son opcionales, lo que permite una gran flexibilidad a la hora de describir a los recursos aun cuando la información que se dispone es parcial o incompleta.

Por tanto, el primer paso ha sido el efectuar un estudio pormenorizado de los repositorios digitales que forman parte de la *Red*, analizando la información

que almacenan y cómo se puede establecer la correspondencia entre estos datos y los elementos básicos de Dublin Core. Este análisis se ha visto dificultado por la ausencia de documentación técnica (como documentos de requisitos o de diseño) de los repositorios, y también por el hecho de que algunos de ellos han sido desarrollados utilizando aplicaciones informáticas propietarias, no abiertas, lo que dificulta el acceso a la información que almacenan.

Como resultado de este análisis se ha concluido que todas las bibliotecas digitales y bases de datos pertenecientes a los grupos que forman parte de la *Red ARACNE* van a poder ser integradas en el *metabusador*. En cualquier caso, para cada una de ellas será preciso diseñar e implementar un sistema mediador (o *proxy*) específico, que permita traducir el formato de datos propio de la biblioteca digital al vocabulario común Dublin Core. La información obtenida en esta fase de análisis resulta de gran utilidad también para depurar los criterios tecnológicos establecidos para la adhesión de una nueva biblioteca digital o base de datos a la *Red*, y que se publicarán en el portal web *Red ARACNE*.

Portal Red ARACNE

En la actualidad se dispone ya de una primera versión pública del portal Red ARACNE, accesible en <http://www.red-aracne.es>. A continuación se describen brevemente las distintas secciones de las que se compone:

1. **Presentación.** Contiene información básica de la *Red*, como la relativa a su motivación original o sus objetivos.
2. **¿Qué ofrecemos?** Ofrece información general sobre lo que puede encontrarse en la *Red ARACNE*, y particular sobre los grupos de investigación, las bibliotecas digitales y las bases de datos especializadas que forman parte de ella.
3. **Publicaciones.** Ofrece una recopilación de las publicaciones relevantes en el ámbito de la *Red*, organizadas en tres secciones: revistas, bibliografía y biblioteca.
4. **Noticias.** Muestra una recopilación de noticias sobre la *Red ARACNE* o cualquier otra información relevante dentro del ámbito de las Humanidades Digitales.
5. **Enlaces.** Ofrece una lista de enlaces a otros espacios web relacionados con el proyecto.
6. **Contacto.** Incluye un formulario que permite contactar con los responsables de la *Red*.

Una sección básica del portal *Red ARACNE*, y que todavía no está disponible, será la interfaz pública del *metabusador*, que permitirá buscar recursos de forma simultánea en todas las bibliotecas digitales y bases de datos integradas en la *Red*.

El portal web *Red ARACNE* incluye una herramienta de administración privada (protegida por contraseña) que permite la gestión de toda la información disponible en él. A esta herramienta de administración pueden acceder tanto los responsables de la gestión de los contenidos de la *Red*, como los responsables de la gestión de los contenidos de cada grupo, si bien estos últimos tendrán acceso restringido únicamente a la información propia de su grupo y de sus bibliotecas o bases de datos.

Por último, destacar que en el desarrollo del portal se ha seguido una arquitectura software modular, a fin de facilitar las más futuras ampliaciones tanto de nuevas secciones como de nuevos grupos y/o bibliotecas digitales³.



RESUMEN

La red-Aracne ofrece una novedad importante: la coordinación de seis proyectos de investigación en Humanidades Digitales (BIESES, BIDISO, BIBLIOTECA SAAVEDRA FAJARDO, CLARISEL, DIALOGYCA BDDH y PHEBO). Se comentan los distintos puntos de partida y los retos que deben abordar para conseguir una coordinación que persigue rentabilizar el trabajo de cada uno de ellos, conseguir más visibilidad en la red y facilitar el acceso de los usuarios a un volumen de información cuantioso y organizado.

Palabras clave: Red de proyectos de investigación, metabuscador.

ABSTRACT

The red-Arachne offers an important novelty: the association of six research projects in Digital Humanities (BIESES, BIDISO, LIBRARY SAAVEDRA FAJARDO, CLARISEL, DIALOGYCA BDDH and PHEBO). We discuss the different starting points and the challenges that must be addressed to achieve coordination in order to optimize the work for each of them, for greater visibility in the web and facilitate users access to a large volume of organized information.

Keywords: Research projects network, Meta searcher.

³ Trabajo realizado en el marco del Proyecto: Inventario, Descripción, Edición Crítica y Análisis de Textos de Prosa Hispánica Bajomedieval y Renacentista. Línea Diálogos (Fase 3) FFI2012-33903 y de la Acción complementaria FFI2011-15606-E (subprograma FILO).