

Ingeniería documental frente a artesanía documental. ¿Cuál es el modelo a seguir?

Por José R. Pérez-Agüera

Resumen: *El presente trabajo se centra en los aspectos presentes y futuros de la profesión, haciendo especial hincapié en la necesidad de modernización de la biblioteconomía y documentación. La propuesta de modernización viene de la mano de las tecnologías informáticas, que permiten perfilar el horizonte de este campo dentro de 20 años.*

Palabras clave: *Biblioteconomía y documentación, Profesión, Futuro de las bibliotecas, Epistemología de la documentación.*

Title: Document engineering vs. artisan documentation: Which model should we follow?

Abstract: *The work is focused on the present and future features of our profession. Especially we are taken into account the modernization needs of Information and Library Science. We propose that this modernization must come from the application of the Computer Science technologies which permit us to draw a new horizon for this area for the next 20 years.*

Keywords: *Library and information science, Profession, Future of libraries, Epistemology of the documentation.*

Pérez-Agüera, José R. "Ingeniería documental frente a artesanía documental. ¿Cuál es el modelo a seguir?". En: *El profesional de la información*, 2008, mayo-junio, v. 17, n. 3, pp. 257-260.

DOI: 10.3145/epi.2008.may.01



José R. Pérez-Agüera es profesor ayudante en el Departamento de ingeniería del software e inteligencia artificial de la Universidad Complutense de Madrid (UCM). Doctor en documentación por la UCM y DEA en ingeniería informática por la misma universidad, donde también cursó la licenciatura en historia, especializándose en historia moderna. Ha realizado estancias de investigación en Yahoo! Research Barcelona y visitado la oficina del W3C del Rutherford-Appleton Laboratory en Oxford. Ha publicado en varios congresos y revistas internacionales y nacionales de las áreas de informática y documentación sobre temas de recuperación de información y web semántica. Es miembro del Consejo asesor de la lista de distribución IweTel y del grupo de trabajo ThinkEPI.

LA INFORMACIÓN ES NUESTRO OBJETO DE ESTUDIO Y DE TRABAJO en todas sus formas, desde los manuscritos del siglo XVII escritos en letra procesal hasta el código fuente que permite a un buscador encontrar los documentos que un usuario necesita. Sin embargo la información tiene una naturaleza muy compleja, y es necesario que los profesionales y científicos de la información sean capaces de investigar y trabajar con su objeto de estudio en todas sus variantes posibles.

El presente de nuestra profesión

Para definir nuestro futuro profesional es necesario entender no sólo nuestro objeto de trabajo sino también su naturaleza. El problema es que en la actualidad esto no es del todo posible. Nuestros alumnos salen bien preparados para enfrentarse a los procesos artesanales de gestión de información que se producen en bibliotecas, archivos y centros de documentación, pero no son

capaces, salvo honrosas excepciones, de implementar un sencillo buscador usando *APIs*¹ ya preestablecidos, o de diseñar un pequeño opac usando *Java*², *MySQL*³ y *Struts*⁴, por poner ejemplos de tecnologías ampliamente extendidas en la creación de sistemas de información.

Por otro lado, si analizamos las principales soluciones empresariales de gestión de información como *Content Manager*⁵, *Documentum*⁶, *Share Point*⁷, *Drupal*⁸, *Joomla*⁹, o incluso el software para la implementación de bibliotecas digitales y depósitos de documentos como *Greenstone*¹⁰ o *Dspace*¹¹, vemos que el uso de lenguajes de programación es imprescindible para la personalización y aplicación de estas herramientas a problemas complejos. Sin embargo nuestros profesionales, en la mayoría de los casos, no son capaces de afrontar este trabajo y deben recurrir a informáticos.

Muchos argumentan que la función de los documentalistas reside en decirle a los informáticos lo que tienen que hacer y que la separación entre "nosotros"

y “ellos” está muy clara. Y puede que esto sea así en la representación platónica de las competencias de profesionales, pero cuando bajamos al mundo real, cuando salimos de la caverna, nos encontramos con que la línea que nos separa de los informáticos se difumina, de forma que éste, mejor o peor, es capaz de hacer el trabajo del documentalista y no al revés. Por eso muchos empresarios, preocupados por el coste de sus proyectos, responden en demasiadas ocasiones contratando únicamente informáticos capaces de afrontar todo el trabajo de principio a fin.

“El informático, mejor o peor, es capaz de hacer el trabajo del documentalista, pero no al revés”

Evidentemente no todos podemos saber de todo. Al igual que un físico puede terminar sus estudios sabiendo mucho de física de materiales y poco o nada de astrofísica, un documentalista debería poder finalizarlos con sólidos conocimientos de programación y poco o nada de paleografía. Pero actualmente esto no es posible en las facultades de documentación de nuestro país, donde todavía hay profesores que se niegan a que la programación, por poner un ejemplo recurrente, sea parte de los temarios de nuestras asignaturas, ya que consideran que no forma parte de las competencias básicas del profesional de la información. Para colmo, el libro blanco de los estudios de documentación en España no puntualiza lo suficiente estos aspectos, dejando realmente a criterio de cada centro la profundidad con la que afrontan este tipo de materias.

Sin duda el problema reside en la forma en la que se ha ido conformando nuestra disciplina en los últimos 30 años, ya que a causa del trasvase de eruditos y humanistas que se produjo en los primeros años a nuestras escuelas y facultades, se ha convertido hoy en una disciplina derivada de las ciencias sociales y las humanidades con una pátina tecnológica bastante laxa.

Uno de los casos más sintomáticos del error de concepto que se ha producido en la formación de profesionales de la información se puede ver en las asignaturas de gestión y recuperación de información que se cursan en muchas de las facultades de nuestro país. La recuperación de información se basa en una serie de algoritmos, que sí se enseñan a los alumnos. Sin embargo no tiene ningún sentido iniciar en algoritmos a gente que no sabe programar. Es como si enseñamos paleografía a alguien que no sabe leer. Sinceramente

no se me ocurre tarea más inútil que instruir sobre el modelo de espacio vectorial a alguien que no es capaz de implementarlo.

Otro ejemplo lo encontramos en asignaturas centradas en el desarrollo de lenguajes documentales, donde cada vez cobran más importancia las ontologías. Es imposible entender qué es una ontología sin tener una base mínima de inteligencia artificial, la cual a su vez necesita de unos conocimientos previos de programación, algoritmia, lógica y estructura de datos, cuestiones todas ellas fuera de los temarios de la mayoría de nuestras facultades.

Los dos casos expuestos son probablemente los más extremos, sin embargo existen otros más: lenguajes de marcado como xml, web semántica, sindicación de contenidos, archivos abiertos (OAI), etc. Todos estos ejemplos revelan un alto grado de inconsistencia entre la formación y lo que se supone que debería ser la práctica profesional ya que, por raro que pueda parecer, enseñar xml a gente sin conocimientos de programación no tiene sentido.

“En la mayoría de los casos, el documentalista no es capaz de comunicarse con el informático en su propio lenguaje”

La principal consecuencia de esta inconsistencia es que pese a existir una amplia demanda de profesionales expertos en nuevas tecnologías, los documentalistas no pueden cubrirla debido a las lagunas que existen en su formación, y esta demanda está siendo cubierta por informáticos, físicos, matemáticos y en general por todo aquel que sepa tirar cuatro líneas de código. Esta situación está produciendo la consiguiente degradación en la definición de procesos de gestión de información donde el documentalista sí tiene mucho que decir.

Algunos dirán que la solución es la conformación de equipos multidisciplinares formados por documentalistas e informáticos, pero el problema reside en que, en la mayoría de los casos, el documentalista no es capaz de comunicarse con el informático en su propio lenguaje, lo que lo convierte en una especie de usuario molesto cuyos requerimientos nunca son totalmente entendidos. Además, muchas veces el “equipo” está compuesto por un único profesional.

Otro ejemplo de la falta de profundidad con la que se tratan los temas tecnológicos en nuestra disciplina y que muestra claramente el problema de la definición profesional, se encuentra en muchos de los artículos de

investigación que se publican en las revistas españolas del ramo y donde encontramos textos divulgativos que pasan como investigación. Desde mi punto de vista, hacer estados del arte o artículos descriptivos sobre tal o cual tecnología no es investigación, y pese a que pueda ser útil en determinados contextos, no puede constituir de ninguna de las maneras el tronco central de las publicaciones de una disciplina.

Sin embargo la gran mayoría de los artículos sobre tecnología en nuestras revistas tienen este enfoque, con la consiguiente escasez de trabajo realmente de investigación. No es que no sea interesante y hasta necesario un artículo sobre Procesamiento de Lenguaje Natural (PLN), pero lo que necesitamos son textos que nos digan cómo se aplica a nuestra disciplina y la forma en la que nos puede ayudar a hacer mejor nuestro trabajo, por ejemplo en la generación automática de tesauros^{8,9}.

Desde la formación a la investigación, pasando por la práctica profesional, encontramos síntomas que nos hablan de los problemas actuales de nuestro campo y que constituyen los principales condicionantes para configurar el futuro de la profesión.

¿Cuál es el futuro?

Es una pregunta imposible de responder, y debo decir con toda sinceridad que no tengo ni idea de la respuesta, pero sí sé cual me gustaría que fuera: que en nuestra profesión se dieran la mano la paleografía y los lenguajes de programación, la catalogación y las bases de datos. Que nuestros alumnos adquieran habilidades como la programación, el diseño de bases de datos –que no es lo mismo que el uso de gestores de bases de datos–, los principios que se encuentran detrás de los lenguajes de marcado, etc. Y todo ello con un conocimiento sólido y profundo del tema, no simplemente que aprendan cuatro tecnicismos de oídas que no llevan a casi nada.

“Se perfila un horizonte muy negro para nuestra profesión si no ampliamos nuestras competencias para hacernos valiosos no sólo hoy, sino también dentro de 50 años”

A la pregunta de cuál es el futuro de nuestra profesión diría que básicamente hay dos posibilidades:

a. Informática documental: si la formación técnica de los documentalistas alcanza niveles aceptables en términos de productividad, el futuro de nuestra profesión estará no sólo en los archivos, bibliotecas y cen-

tros de documentación tradicionales como los de los medios de comunicación, sino también en el mundo de la empresa ayudando a rentabilizar y capitalizar su recurso más valioso: la información.

b. Artesanía documental: si las cosas siguen como están ahora, el futuro se reducirá casi exclusivamente a bibliotecas, archivos y lugares donde la informatización de la información no sea posible. El problema de este futuro reside en que si tenemos en cuenta el avance tecnológico previsto para los próximos años, el nicho para los documentalistas se irá reduciendo progresivamente hasta desaparecer. Se perfila un horizonte muy negro para nuestra profesión si no ampliamos nuestras competencias hasta el punto no sólo de hacernos valiosos hoy, sino que sigamos siéndolo dentro de 50 años. En pocas palabras: que evolucionemos.

En mi opinión la mayoría de los argumentos en favor del establecimiento de límites entre informática y documentación no sirven sino para ocultar la falta de ganas de reciclarse que tienen muchos de los profesionales y profesores de nuestra área. Creo que nadie debe autoproclamarse experto en tecnologías de la información si no sabe programar, si nunca ha diseñado e implementado un sistema de información. Es como considerarse arquitecto sin tener la más remota idea de cómo se calcula la estructura de una casa.

“No podemos pedirles a los informáticos que no sólo nos hagan la carretera sino que además nos lleven a cuestas a lo largo de ella”

Desde mi punto de vista, la ingeniería documental viene definida por una serie de habilidades y competencias necesarias para afrontar con suficientes garantías la implantación de sistemas complejos de gestión de información. A grandes rasgos, estas competencias estarían relacionadas con la adquisición de los siguientes conocimientos:

- Teóricos y prácticos de programación, incluyendo algoritmia y estructuras de datos.
- Teóricos y prácticos de ingeniería del software.
- Teóricos y prácticos de sistemas de ficheros y bases de datos relacionales.
- Básicos de gestión y administración de sistemas informáticos.
- De inteligencia artificial, tanto enfoques simbólicos como subsimbólicos.

Se puede argumentar que si estas son las competencias que definen a un documentalista técnico o ingeniero documental ¿qué le diferencia entonces de un ingeniero informático? La respuesta es sencilla, un informático no sólo sabe programar, puede diseñar lenguajes de programación, un informático no sólo es capaz de administrar un servidor, tiene conocimientos sobre sistemas operativos que les permiten implementar la infraestructura sobre la que se instala. En definitiva, construye las carreteras por las que los profesionales de la información debemos transitar, nos aportan la infraestructura sobre la que construir los sistemas de información y diseñan las herramientas que nos permiten ejecutar esta tarea, pero lo que no podemos ni debemos pedirles es que no sólo nos hagan la carretera sino que además nos lleven a cuestas a lo largo de ella.

“En la sociedad de la información en la que vivimos las oportunidades para nuestra profesión van mucho más lejos”

Un documentalista técnico, un ingeniero de la información, no es necesario que sepa elaborar sistemas operativos, pero sí que conozca en detalle como funciona el software que hace que un ordenador sea algo más que un pisapapeles de 6.000 €. Tampoco ha de ser capaz de llevar a cabo un compilador, pero sí es imprescindible que domine los lenguajes de programación más extendidos y no solamente pueda proponer ideas, sino también implementarlas aunque sea en la forma de un simple prototipo.

Existe una necesidad imperiosa de que el documentalista atraviese el interfaz de las aplicaciones, vaya más allá y se inmiscuya en el desarrollo de las mismas.

Sin duda el futuro de la profesión está garantizado desde el punto de vista de la artesanía documental, al menos para los próximos 20 años. Bibliotecas y archivos parecen de momento cotos privados y reservados para nuestra profesión y seguramente lo seguirán siendo durante bastante tiempo, aunque posiblemente no de forma indefinida. Ahora bien, en la sociedad de la información en la que vivimos las opciones de desarrollo profesional para nuestra profesión van mucho más lejos.

La decisión de aprovechar esta oportunidad depende de nosotros, y supone un enorme esfuerzo por nuestra parte ya que debemos alejarnos, aunque sea un poco, de las ciencias sociales y las humanidades para acercarnos a las matemáticas y a las ciencias de la computación. Muchos dirán que este viaje ya lo hemos hecho, pero yo creo que todavía no se ha empezado y

“Debemos alejarnos, aunque sea un poco, de las ciencias sociales y las humanidades para acercarnos a las matemáticas y a las ciencias de la computación”

que si no avanzamos bastante en este sentido perderemos una oportunidad única de renovación profesional y nos quedaremos quejándonos periódicamente en listas como *IweTel* de lo poco que se nos valora. Y lo que es más importante, si no damos a la sociedad lo que nos demanda, nuestra profesión desaparecerá como lo han hecho muchas otras.

Notas y referencias

1. Una *API* (del inglés *Application programming interface*-Interfaz de programación de aplicaciones) es el conjunto de funciones y procedimientos (o métodos si se refiere a programación orientada a objetos) que ofrece cierta biblioteca para ser utilizado por otro software como una capa de abstracción. Definición tomada de la *Wikipedia*.
http://es.wikipedia.org/wiki/Application_Programming_Interface
2. *Java* es un lenguaje de programación orientado a objetos desarrollado por *Sun Microsystems* a principios de los años 90. Definición tomada de la *Wikipedia*.
http://es.wikipedia.org/wiki/Lenguaje_de_programaci%C3%B3n_Java
3. *MySQL* es un sistema de gestión de base de datos relacional, multihilo y multiusuario con más de seis millones de instalaciones. Definición tomada de la *Wikipedia*.
<http://es.wikipedia.org/wiki/MySQL>
4. *Struts* es una herramienta de soporte para realizar aplicaciones web bajo el patrón *MVC* bajo la plataforma *J2EE* (*Java 2, Enterprise Edition*). Se desarrollaba como parte del proyecto *Jakarta* de la *Apache Software Foundation*, pero actualmente es un proyecto independiente conocido como *Apache Struts*. Definición tomada de la *Wikipedia*.
http://es.wikipedia.org/wiki/Apache_Struts
5. <http://www-306.ibm.com/software/data/cm/docmgr/>
6. <http://www.documentum.com/>
7. <http://office.microsoft.com/es-es/sharepointserver/FX100492003082.aspx>
8. <http://www.joomlaspanish.org/>
9. <http://www.drupal.org.es/>
10. <http://www.greenstone.org/>
11. <http://www.dspace.org/>
12. Pérez-Agüera, José R.; Araujo, Lourdes. «Query expansion with an automatically generated thesaurus». Corchado, E., et al. (eds.). *Lecture notes in computer science*, 2006, n. 4.224, pp. 771-778.
13. Araujo, Lourdes; Pérez-Agüera, José R. «Enriching thesauri with hierarchical relationships by pattern matching in dictionaries». En: Salakoski, T., et al. (eds.). *Lecture notes in artificial intelligence*, 2006, n. 4.139, pp. 268-279.

José R. Pérez-Agüera, Departamento de ingeniería del software e inteligencia artificial, Universidad Complutense de Madrid.
jose.aguera@fdi.ucm.es