Aplicación de las tecnologías de la información a la investigación en Historia

Vivina Asensi Artiga
Pedro Manuel Díaz Ortuño
Francisco Javier Martínez Méndez
Gregorio Moya Martínez
José Vicente Rodríguez Muñoz
Profesores de la E. U. de Biblioteconomía y Documentación
Universidad de Murcia

Las Tecnologías de la Información permiten el tratamiento documental de las fuentes históricas mediante la aplicación de programas standard, que posibilitan el intercambio de información entre los investigadores. El uso de software específico para cada una de las facetas de la historiografía, o incluso para cada uno de los estadios de un mismo proyecto, proporcionan óptimos resultados. Metodología empleada en el tratamiento de documentación correspondiente a finales del siglo XV.

Information Technologies permit the treatment of the information present in historical sources through the application of standard programmes in order to facilitate the exchange of information among researchers. The use of specific software in each of the facets of history science or even in each of the stages of the same project yields optimal results. Methodology used in the treatment of documents dating back to the end of the fifteenth century.

1. Introducción

El estudio de las fuentes históricas, máxime las pertenecientes a la Edad Media y también a la Edad Moderna, requiere una evaluación detenida de la información que se quiere conservar y de la que se puede prescindir. La década de los ochenta se caracterizó por la concienciación de los historiadores acerca de la necesidad de trabajar en equipo para el tratamiento documental de dichas fuentes, mediante las Tecnologías de la Información. En Gottingen (1985), Graz (1986) y en París (1987), se celebraron reuniones y jornadas de estudio con la finalidad de constituir grupos de trabajo para estudiar los problemas presentados por la normalización e intercambio de bases de datos sobre temas de historia[1].

Es evidente que las características peculiares de las humanidades y de las ciencias sociales plantean la necesidad de almacenamiento y recuperación de la in-

1. GENET, J. PH., "Introduction" En: *Standardisation et échange des bases de données historiques : actes de la troisième Table Ronde internationale...*, Paris, 1987, p. 15-21

Esta colección documental puede confeccionarse mediante la aplicación de un editor de textos, siguiendo el formato acostumbrado: número de documento en la colección, fecha, resumen de contenido, fuente y transcripción literal del texto completo.

El procesador de textos elegido ha de ser potente y de fácil manejo, condiciones necesarias para trabajar de una forma eficaz, proporcionándonos gran cantidad de utilidades que permitan un trabajo de edición más cómodo y llevadero. De los distintos programas de edición de texto de los que disponíamos nos decantamos hacia Wordperfect (WP) en su versión 5.1, por la potencia y versatilidad que proporciona. Además, dicho procesador ofrece una serie de utilidades muy adecuadas para ejecutar las tareas que conlleva la edición del texto, entre las que destacamos especialmente dos: la utilidad de Documento Maestro y la de generación automática de índices y tablas de contenido, de las que hablaremos posteriormente.

Si bien no se trata aquí de ensalzar las virtudes de dicho programa de edición, ni de hacerle una propaganda gratuita, si que nos creemos en la obligación de proporcionar una breve información sobre las características que posee WP 5.1 que nos han sido de mucha utilidad, deteniéndonos un poco más en las dos mencionadas anteriormente.

WP 5.1 es un programa de tratamiento de textos avanzado con posibilidades de autoedición. De entre sus muchas características nos gustaría destacar las siguientes:

- Permite realizar de una manera cómoda y sencilla las tareas típicas de procesamiento de texto: tecleo de texto, movimientos de bloques, márgenes, tabulaciones y todas aquellas operaciones de formato del texto.

- Permite escribir el texto encolumnado, verificar la ortografía para encontrar posibles errores ortográficos y obtener sugerencias sobre el uso de palabras alternativas, gracias al diccionario de sinónimos que incorpora.

- Se puede cambiar entre los distintos tipos y tamaños de letra con gran facilidad. Cabe mencionar en este punto que la variedad de tipos de letra dependerá en gran medida de la impresora de la que dispongamos, de igual manera que la calidad de los textos impresos. Nosotros tuvimos la posibilidad de trabajar con una impresora láser con una amplia gama de tipos de letra, con lo que la calidad de los textos impresos era cercana a la de una imprenta.

- La realización de notas a pie de página es sencilla, ya que solo debemos escribir el texto que queremos incluir en la nota, y el programa se encarga de dejar el espacio necesario al final de la página, para imprimir la o las notas deseadas.

- Otra de las utilidades a considerar es la protección de bloques separados, mediante la cual, el programa ajusta el texto de manera que no quede una línea aislada del resto del párrafo al final (o al principio) de una página.

- También dispone del efecto WYSIWYG (What You See Is What You Get), que nos permite "ver", mediante una representación gráfica en la pantalla, cómo nos quedaría el texto una vez impreso, con el ahorro en papel que esto conlleva al poder suprimir casi por completo las pruebas de impresión.

Pero sobre todo, las utilidades que más nos ayudaron en la realización del texto fueron las ya mencionadas anteriormente: la de Documento Maestro, y la de generación automática de índices y tablas de contenido.

- La utilidad de Documento Maestro nos permitió escribir nuestro trabajo divi-

diéndolo en pequeños ficheros, lo cual hacía más manejable la edición y manipulación del texto, además de aumentar la seguridad de nuestro trabajo. (Es más fácil trabajar con un fichero que contenga unas pocas páginas de texto, que con uno grande; además, si por alguna causa perdemos dicho fichero, la cantidad de trabajo que tendríamos que rehacer es mucho menor).

Mediante esta utilidad, uniendo todos los ficheros en uno solo (cuando se ha terminado de introducir todo el texto), el programa gestiona de forma automática la numeración de las páginas, los códigos de formatos, las marcas de los índices, etc., lo cual permite escribir los códigos de formato del documento una sola vez, con lo que se consigue una presentación más uniforme del documento.

- La utilidad de generación de índices y tablas de contenido resulta de gran ayuda a la hora de realizar dichos índices, ya que sólo es necesario marcar el texto que se considere oportuno y especificarle en qué índice deseamos incluirlo. (Nosotros realizamos tres índices: onomástico, de materias y de topónimos. De esta forma pudimos realizar los tres índices a la vez). Al finalizar el proceso de marcar el texto que debe contener el índice (se puede realizar al mismo tiempo que se teclea el texto, aunque las marcas estén en ficheros separados), sólo es necesario definir el formato que queramos dar a nuestros índices e indicar al programa que deseamos generarlos. De este modo, los índices solicitados se realizan automáticamente en breves momentos.

Hartel[7] considera responsabilidad del historiador la decisión correspondiente a la explicación de los términos, máxime cuando se adjunte un glosario a la edición de los documentos, hasta tal punto que el término seleccionado pueda, o no, pertenecer a la fuente. A este respecto, consideramos incluir los términos correspondientes a los índices onomástico y toponímico respetando los términos originales; aunque, con los reenvíos correspondientes, es decir con referencias que remitan de la forma actual a la forma utilizada en la fuente. Utilizamos referencias cruzadas para las distintas formas que presenta un mismo término. No sucede lo mismo con el índice de materias, cuyas entradas corresponden al lenguaje actual.

3.2. Creación de la estructura de la base de datos

Ateniéndonos a las características propias de la historiografía actual, los Sistemas de Gestión Documental son adecuados para tratamiento óptimo de la documentación con miras a una recuperación rápida, fiable y pertinente. Por ello descartamos Sistemas de Gestión de Bases de Datos convencionales, como por ejemplo las Bases de Datos Relacionales, más apropiados para procesar valores numéricos o cadenas alfanuméricas relativamente simples.

Se seleccionó el programa Knosys, y no otros -como Freebase, Askam, Micro-Questel, BRS-Search o BASIS- por las siguientes características: facilidad de uso, agradable interface con el usuario y potencia de su lenguaje de interrogación.

La estructura de la base de datos se creó respetando las normas catalográficas utilizadas generalmente en los archivos históricos: fecha, signatura, tipología docu-

7. HARTEL, R., "Le texte suffit-il?" En: *Standardisation...*, p. 33-48

mental y asunto; aunque resulta conveniente añadir otros dos campos: descriptores y resumen.

Los descriptores pueden coincidir con las entradas del índice de materias de la colección documental, que contiene el texto integro del documento. En este campo, aunque sigue utilizándose el lenguaje natural, es evidente cierto control sobre la terminología. El diccionario de sinónimos, por otra parte, permite ampliar las posibilidades de recuperación.

Pese a que el logicial permite la recuperación full-text, la terminología utilizada no cuenta con la uniformidad indispensable para una correcta indización. Los truncamientos necesarios para la recuperación serían excesivos y la respuesta a la búsqueda realizada resultaría tan exhaustiva, que muy bien podría confundirse con ruido documental. En consecuencia es aconsejable introducir la información en lenguaje actual.

No es conveniente la interpretación subjetiva de los documentos históricos. La Paleografía requiere repetidas lecturas y, en ocasiones, el perfil de investigación puede condicionar las conclusiones obtenidas. En consecuencia, nos parece más adecuado introducir un breve resumen del contenido del documento para que el investigador, o sencillamente el estudioso con inquietudes de conocer determinadas facetas de la Historia, examine las fuentes con el espíritu critico necesario.

La signatura nos remitirá tanto al corpus documental como a la fuente original.

3.3. Posibilidades de importación y exportación

El SGD Knosys permite la importación de ficheros Ascii para su almacenamiento en una base de datos. El programa posee un editor de textos integrado; no obstante los documentos de la base de datos pueden exportarse a un programa de edición de textos más completo, por ejemplo WP 5.0 ó WP 5.1, generando un fichero en disco.

3.4. Posibilidad de ampliar las aplicaciones del software

Los lenguajes de recuperación, por sencillos que parezcan en cuanto se conocen, suponen un obstáculo para quienes se enfrentan a ellos por primera vez, hasta el punto de ser uno de los factores que influyen de manera decisiva en el rechazo de algunos investigadores, acostumbrados al proceso documental convencional mediante técnicas manuales, ante las Tecnologías de la Información.

El thesaurus permitiría realizar las consultas en la base de datos mediante relaciones, jerárquicas o asociativas, entre los términos por los que se perfila la búsqueda; aunque no debemos ignorar el riesgo de generar ruido documental.

En el marco de trabajo del SGD Knosys, no existe la posibilidad de introducir un thesaurus para paliar los inconvenientes que presenta el lenguaje natural para la recuperación[8]. Por ello apuntamos

8. KRISTENSEN, J. y JARVELIN, K., "The effectiveness of a searching...", p. 77-78

la posibilidad de crear un thesaurus fuera del entorno Knosys en un módulo independiente[9].

El proceso a desarrollar para llevar a cabo esta aplicación deberá seguir las normas propias del análisis documental y, en consecuencia, el número de descriptores asociado a cada documento será reducido. De este modo podremos crear nuestro propio thesaurus sin un coste de programación elevado, siempre y cuando poseamos ciertos conocimientos de Lenguajes de Programación.

Otra posibilidad, hacia la que dirigimos nuestra investigación consiste en la implantación de un sistema de mantenimiento de thesaurus para gestores documentales bajo la perspectiva del hipertexto, debido a las facilidades que proporciona al respecto.

Por último, debemos mencionar la facilidad con la que pudimos pasar los ficheros que contenían los documentos tratados con el SGD Knosys a WP, para que la impresión de dichos documentos fuera coherente con el resto del texto.

4. Conclusiones

Las Tecnologías de la Información ponen nuevas herramientas al servicio del historiador. El uso de programas standard permite mayor eficacia en la difusión de los resultados obtenidos en la investigación y, en consecuencia, la posibilidad real de que puedan servir de punto de partida para trabajos posteriores, metas siempre deseadas pero difíciles de conseguir por falta de medios.

El tipo de software adecuado está en función del tratamiento que se desee aplicar.

Los procesadores de texto permiten la edición de colecciones documentales, proporcionan gran facilidad y versatilidad en todas sus aplicaciones: edición de los documentos; generación y edición automática de índices y tablas de contenidos; generación de gráficos, estadísticas... si se desea, además de proporcionar una gran variedad de formatos y estilos para la impresión del texto.

Los SGD permiten la creación de bases de datos en formato de longitud variable y la posibilidad de recuperación de la información en texto completo, por lo que son adecuados para la documentación en Ciencias Sociales y Humanidades. Algunos programas resultan especialmente apropiados para el historiador por la facilidad que ofrece el interface con el usuario; además la posibilidad de utilizar el lenguaje natural en la indización, atrae a los investigadores no expertos en técnicas documentales.

La capacidad de importar-exportar datos, así como futuras prestaciones que se pueden obtener por la conexión con otros programas, o sencillamente por la aplicación de otras técnicas o innovaciones que ofrecen las Tecnologías de la Información, constituyen argumento suficiente para tener en cuenta el tratamiento documental de la documentación en Humanidades y Ciencias Sociales, y concretamente de las fuentes históricas, con ayuda de programas informáticos.

9. ASENSI ARTIGA, V. y otros, "Implementación de un módulo de gestión de tesauros en un Sistema Gestor Documental" En: *4es Jornades Catalanes de Documentació,* Barcelona, SOCADI, 1992, p. 361