

Informe sobre avances en la informatización del Archivo Rubén Darío (www.ucm.es/info/rdario)

Niall BINNS, Carlos FERNÁNDEZ GONZÁLEZ y Olga MUÑOZ CARRASCO

Universidad Complutense de Madrid

RESUMEN

Este breve texto ofrece información sobre los avances que se han hecho en la informatización de los casi cinco mil documentos del Archivo Rubén Darío. Presenta la página web del Archivo, comenta las dificultades experimentadas en la labor de informatización, ofrece detalles técnicos sobre la digitalización y transcribe un índice de los contenidos del Archivo para facilitar la labor del investigador.

Palabras clave: Rubén Darío, Archivo

Report on progress in the computerisation of the Archive Rubén Darío

ABSTRACT

This brief text offers information on the progress that has been made in the computerisation of the nearly five thousand documents of the Archive Rubén Darío. It presents the Archive's web page, comments some of the difficulties that have been faced, offers technical details on the computerisation and an index of the Archive's contents so as to enable researchers to travel through the web page.

Key words: Rubén Darío, Archive

SUMARIO: 1. La digitalización del Archivo Rubén Darío. 2. Índice de las carpetas que constituyen el Archivo «Rubén Darío»

En el número 32 de *Anales de Literatura Hispanoamericana* (2003), el profesor Luis Sáinz de Medrano Arce, presidente de la Comisión Seminario-Archivo Rubén Darío, ofreció una breve historia sobre el Archivo Rubén Darío (reproducida en la página web indicada arriba). Como se señaló entonces, bajo la dirección del Dr. Sáinz de Medrano y con la coordinación del Dr. Niall Binns, los investigadores Olga Muñoz Carrasco, Carlos Fernández González y José Manuel González Freire han estado trabajando en la informatización de las 79 carpetas y los 4.795 documentos que constituyen el Archivo. Esta labor, que se ha hecho posible gracias a los avances tecnológicos de la última década, resulta particularmente importante dado el estado de los documentos —sometidos al desgaste del tiempo y del uso— y su evidente valor —el riesgo de «pérdidas» voluntarias o involuntarias (en las últimas décadas se han «traspapelado», por ejemplo, documentos de Leopoldo Lugones y Delmira Agustini)—, y consiste fundamental-

mente en lo siguiente: escanear los documentos; transcribirlos; introducir los documentos escaneados y transcritos en una base de datos; y abrirlos a la consulta pública a través de Internet, para que puedan tener acceso a ellos no sólo los especialistas —sin necesidad de trasladarse a Madrid—, sino también los estudiantes del modernismo y el público general.

Los avances importantes que se han hecho en los últimos años se deben, en gran medida, a dos «Acciones Especiales» otorgadas por el Ministerio español de Ciencia y Tecnología, y en este momento 3.906 imágenes han sido incorporadas en la página web (en fase de pruebas), además de las transcripciones de aproximadamente la mitad de los documentos del Archivo.

Conviene señalar toda una serie de problemas u obstáculos con los que hemos tenido que lidiar: la vastedad del Archivo; el analfabetismo informático habitual en los filólogos (afortunadamente, hemos podido contar con Carlos Fernández González y José Manuel González Freire); la necesidad de un control minucioso del Archivo y una consiguiente restricción en el personal integrado a la labor de informatización de los originales; la extensión de muchos de los documentos (la transcripción y la digitalización de documentos que consisten en varias hojas); el mal estado y/o la ilegibilidad de muchos documentos, con la evidente dificultad que esto implica para la transcripción; las interrupciones exasperantes provocadas por la tramitación de las solicitudes de financiación: entre la primera y la segunda Acciones Especiales, hubo que esperar ocho meses, y a mediados de abril de 2005 se nos ha desestimado, por motivos formales, la tercera ayuda económica solicitada al Ministerio en septiembre de 2004.

Como se podrá ver al entrar en la página web (www.ucm.es/info/rdario), ésta sigue en fase de pruebas. Hemos querido dar prioridad, por ahora, a la cantidad de documentos accesibles a través de la red que a la calidad de su presentación. Queda, todavía, mucho por hacer. Hace falta: digitalizar y transcribir más de 1.500 documentos; elaborar un índice del catálogo (al final de este breve informe se ofrece una pequeña guía para el investigador); emprender una corrección meticulosa de las transcripciones (hasta el momento, éstas se han hecho a partir de fotocopias, y han sido incorporadas en la base de datos sin corregir, con errores fácilmente remediabiles, y a menudo con puntos suspensivos o puntos de interrogación para indicar dificultades de comprensión; por otro lado, habrá que comparar las transcripciones con las que ya se han hecho en los diversos epistolarios ya publicados por especialistas que han consultado los fondos del Archivo); facilitar las búsquedas; rellenar las distintas casillas de la base de datos con información sobre los documentos; asegurar una difusión más amplia de la página web.

La página web incluye una página de presentación (el artículo ya citado del Dr. Sáinz de Medrano); un documento sonoro correspondiente a la entrevista, grabada por la profesora norteamericana Evelyn Uhrhan, que le hizo a Francisca Sánchez el profesor Antonio Oliver Belmás en 1957; y una página de búsqueda donde se puede acceder a todos los documentos introducidos en la base de datos.

LA DIGITALIZACIÓN DEL ARCHIVO RUBÉN DARÍO

La digitalización mediante escáner de los documentos se ha efectuado originalmente en formato PDF. Para su integración en la web, los documentos se han convertido al formato JPEG y se ha procedido a una posterior optimización. El resultado final ha sido la obtención de 3.906 imágenes. Mientras tanto, los documentos transcritos en word han sido luego insertadas en ACCESS en la base de datos normalizada.

BASE DE DATOS EN ACCESS XP

En primer lugar, se han establecido los criterios de normalización para el Archivo Rubén Darío. En esta fase se ha normalizado el proceso de obtención de datos para el establecimiento de un modelo relacional de base de datos en el cual se procesa la información para el desarrollo de posteriores trabajos de inserción de datos de campo (transcripciones, imágenes asociadas a documento, manipulación y transformación de ficheros intermedios y generación de contenido estático para la web).

Como resultado del modelo relacional se ha generado una aplicación en VBA (Visual Basic para Aplicaciones), facilitando la introducción de datos de forma sencilla y normalizada. Igualmente existe un módulo basado en registro relacional que genera páginas web estáticas normalizadas con los contenidos derivados de la obtención de los datos del Archivo. También permite integrar en un entorno navegacional la inserción de imágenes facsímiles derivadas de procesos de escaneo TWAIN.

La estructura de base de datos a través de un módulo de extracción de recursos de imagen normaliza la inserción de imágenes estableciendo un sistema de nomenclatura de ficheros que permite identificar la pertenencia, la cantidad y la navegabilidad del producto final obtenido en HTML estático garantizando la dualidad de registros y del contenido de las páginas web.

La estructura de directorios de la aplicación base en VBA garantiza que la normalización de la base de datos y los resultados que finalmente se visualizan vía Internet son los correctos, teniendo que realizarse la fase de transferencia de ficheros de forma manual para garantizar la actualización concreta de aquellos datos de nueva aparición.

HTML

El módulo de generación de resultados HTML sobrescribe la estructura de la web completa, siendo una persona la encargada de actualizar en fase de transferencia de ficheros aquellas carpetas que corresponden a modificaciones de registro en la base de datos. En todo caso es responsabilidad última de la persona encargada la actualización de contenidos basándose en el mantenimiento de un

histórico de actualizaciones que conformarán el registro de contenidos actualmente colgado en Internet.

TAMAÑO DE CARPETAS DE SERVIDOR FTP Y SERVIDOR WEB

Se ha hecho un esfuerzo importante entre el compromiso de calidad y tamaño de ficheros, principalmente de imágenes, para no saturar los tiempos de carga ni fomentar el crecimiento exponencial de recursos de servidor, por lo que en el momento actual se establece en un tamaño en disco aproximado de 312 MB encontrándonos en el 65% del proceso de digitalización y obtención de datos normalizados del Archivo. Así pues creemos que las 600 MB ofrecidas por ustedes deben de ser suficientes para completar este proyecto y depurar finalmente el resultado pretendido.

MOTORES DE BÚSQUEDA

Se ha normalizado la generación de metatags de archivo HTML obteniendo aquellos datos que son relevantes y omitiendo las entradas no válidas necesarias en la generación de documento estático a partir de registro de la base de datos del Archivo.

Este punto es importante porque la base de datos no está completa y en el futuro queda pendiente la inserción de metatags keywords que corresponderán a descriptores de documentos no introducidos en la actualidad. Se restringen así las búsquedas a aquellos documentos en los cuales sí se han actualizado los correspondientes descriptores.

Para facilitar la obtención de datos vía Internet mediante buscador se ha utilizado un motor de búsqueda basado en Javascript 1.2, puesto que no se han implementado tecnología de acceso a datos contra bases de datos de servidor.

Se han generado tres motores de búsqueda atendiendo a criterios de Metatag:

Keywords: correspondiente con los descriptores definidos para cada documento según el contenido de registro de la base de datos del Archivo.

Texto completo: correspondiente con el contenido de todos los campos de cada página web generada según el contenido de registro de la base de datos del Archivo.

Título (title): correspondiente con la signatura del documento que identifica de forma unívoca a cada página web.

La aplicación base utilizada para la generación de los scripts ha sido el Search Engine Builder Professional v.1.7.7 registrada a nombre del Seminario-Archivo Rubén Darío.

Esta aplicación es de utilidad para generar nuevos motores de búsqueda a medida que se completa la actualización de datos del archivo.

Se han configurado los generadores de motores de búsqueda en base a los anteriores criterios de metatags para facilitar su actualización vía FTP en el caso

de una actualización significativa de contenidos de la base de datos.

AUTOMATIZACIÓN EN LA OPTIMIZACIÓN DE IMÁGENES NORMALIZADAS

Se ha utilizado la aplicación base Arles Web Page Images Creator, para facilitar la homogeneización del tamaño en píxeles de todos los documentos gráficos del Archivo y su correcta visualización al incrustarse en los documentos HTML generados por la aplicación VBA ACCESS desarrollada al efecto.

ÍNDICE DE LAS CARPETAS QUE CONSTITUYEN EL ARCHIVO «RUBÉN DARÍO»

- Carpeta 1. Libro de copias nº. 1 (cartas de Rubén Darío): documentos 1-146.
- Carpeta 2. Libro de copias nº. 2 (cartas de Darío): 147-409.
- Carpeta 3. Testamentos de Darío: 410-412.
- Carpeta 4. Epistolario íntimo con Francisca Sánchez: 413-469.
- Carpeta 5. Otros autógrafos: 470-532; 4786-4788.
- Carpeta 6. Poemas autógrafos: 533-543.
- Carpeta 7. Fotografías: 544-630.
- Carpeta 8. Cartas a Darío desde Argentina: 631-704.
- Carpeta 9. Cartas desde Argentina: 706-811.
- Carpeta 10. Cartas desde Bolivia: 827-860.
- Carpeta 11. Cartas desde Brasil: 863-868.
- Carpeta 12. Cartas desde Colombia: 869-933.
- Carpeta 13. Cartas desde Costa Rica: 950-963.
- Carpeta 14. Cartas desde Cuba: 964-1017.
- Carpeta 15. Cartas desde Chile: 1018-1054.
- Carpeta 16. Cartas desde Ecuador: 1063-1071.
- Carpeta 17. Cartas desde El Salvador: 1072-1081.
- Carpeta 18. Cartas desde Guatemala: 1082-1133.
- Carpeta 19. Cartas desde Haití: 1134.
- Carpeta 20. Cartas desde Honduras: 1135-1140; 4616-4617.
- Carpeta 21. Cartas desde México: 1141-1184.
- Carpeta 22. Cartas desde México: 1185-1287; 4770.
- Carpeta 23. Cartas desde Nicaragua: 1288-1338; 2185-2190.
- Carpeta 24. Cartas desde Nicaragua: 2191; 1339-1433.
- Carpeta 25. Cartas desde Panamá: 1434-1446.
- Carpeta 26. Cartas desde Paraguay: 1447-1449.
- Carpeta 27. Cartas desde Perú: 1450-1483.
- Carpeta 28. Cartas desde Puerto Rico: 1484-1524; 4777-4778.
- Carpeta 29. Cartas desde Santo Domingo: 1525-1593.
- Carpeta 30. Cartas desde Uruguay: 1594-1651.

- Carpeta 31. Cartas desde Venezuela: 1652-1701.
Carpeta 32. Cartas desde España: 1702-1768; 2072; 4525-4589.
Carpeta 33. Cartas desde España: 1769-1879; 2067-2070; 4549-4553.
Carpeta 34. Cartas desde España: 1880-1987.
Carpeta 35. Cartas desde España: 1988; 2066; 4554-4565.
Carpeta 36. Cartas desde Alemania: 2073-2076.
Carpeta 37. Cartas desde Austria: 2077.
Carpeta 38. Cartas desde Dinamarca: 2078-2080; 4523.
Carpeta 39. Cartas desde Bélgica.
Carpeta 40. Cartas desde Francia: 2081-2151; 4772.
Carpeta 41. Cartas desde Inglaterra: 2135-2136; 4522.
Carpeta 42. Cartas desde Italia: 2137-2144; 4515-4521.
Carpeta 43. Cartas desde Portugal: 2145-2146.
Carpeta 44. Cartas desde Turquía: 2147.
Carpeta 45. Cartas desde Filipinas.
Carpeta 46. Cartas desde Estados Unidos: 2148-2149; 2373; 4771; 4779.
Carpeta 47. Cartas a Francisca Sánchez - Pésames: 2152-2184; 2659.
Carpeta 48. Otros documentos en relación con Francisca Sánchez: 2192-2234; 4782-4785.
Carpeta 49. Documentos relativos a Rubén Darío Sánchez (Güicho): 2236-2260.
Carpeta 50. Petición de autógrafos: 2261-2385.
Carpeta 51. Homenajes a Rubén Darío: 2332-2386.
Carpeta 52. Carta de Rubén Darío Contreras: 2381.
Carpeta 53. Cartas de admiradores: 2387-2399.
Carpeta 54. Traducción de obras de Darío: 2403.
Carpeta 55. Documentos de Historia Contemporánea: 2404-2479.
Carpeta 56. Ediciones rubenianas: 2480-2525; 2906-2908.
Carpeta 57. Editoriales y revistas: 2526-2688.
Carpeta 58. Editoriales y revistas: 2689-2901.
Carpeta 59. Editoriales y revistas: 2902-2947.
Carpeta 60. Diarios, revistas y recortes de prensa: 2948-3019.
Carpeta 61. Diarios, revistas y recortes de prensa: 3020-3084; 4781; 4790.
Carpeta 62. Tarjetas de vista: 3085-3447.
Carpeta 63. Tarjetas de Rubén Darío: 3448-3540.
Carpeta 64. Originales diversos: 3541-3712.
Carpeta 65. Carrera diplomática: 3713-3767; 4773.
Carpeta 66. Carrera diplomática: 3768-3878.
Carpeta 67. Carrera diplomática: 3879-4010.
Carpeta 68. Facturas y cuentas: 4011-4180.
Carpeta 69. Facturas y cuentas: 4181-4285.
Carpeta 70. Argentina: 4286-4328; 4443-4444; 4595; 4600-4606; 4789.
Carpeta 71. Bolivia, Brasil, Colombia, Cuba: 4329-4337; 4446; 4607-4610.
Carpeta 72. Colombia, Costa Rica, Cuba, Chile, Ecuador, Guatemala, Honduras: 4338-4366; 4441-4442; 4424; 4612-4615.

Carpeta 73. Ecuador, El Salvador, Guatemala: 4367-4382; 4425; 4468-4469.

Carpeta 74. México, Nicaragua: 4383-4448; 4618-4651; 4774-4776.

Carpeta 75. Perú, Santo Domingo, Uruguay, Venezuela: 4426-4440; 4652-4667.

Carpeta 76. Tratamientos médicos: 4449-4465.

Carpeta 77. Rusia: 4524.

Carpeta 78. Alemania, Francia: 4466-4514.

Carpeta 79. Nicaragua: 4780-4795.