

INGENIERÍA DOCUMENTAL ORIENTADA A OBJETOS: UN MÉTODO DE DESARROLLO DE SISTEMAS DE INFORMACIÓN DOCUMENTAL

1 Introducción

El término Ingeniería Documental hace referencia al conjunto de métodos, técnicas y herramientas para el diseño, construcción e implantación de sistemas de información documental. La conveniencia de abordar el desarrollo de estos sistemas como un proceso de ingeniería se justifica cuando la complejidad de los documentos implicados y de su tratamiento exige la aplicación de un método sistemático y riguroso que garantice un adecuado nivel de calidad. Éste sería el caso, por ejemplo, del desarrollo de documentación hipermedia en la World Wide Web, imposible de mantener si no se realiza previamente un diseño de la estructura de los documentos (1). Otro caso que también requiere del rigor de un método de ingeniería es la producción y explotación de la documentación utilizada en el proceso de negocio de una organización, ya que, en este caso, es evidente la importancia de los documentos para la organización, al contener la información necesaria para el correcto desarrollo de la actividad de la misma (2).

La Ingeniería Documental es una disciplina relativamente reciente (las primeras referencias son de principios de los años 90). Actualmente existe una gran actividad investigadora en este campo, especialmente relacionada con el establecimiento de metodologías de desarrollo de sistemas de información documental, y con la construcción de herramientas informáticas que faciliten la realización de las diferentes actividades documentales. El término que suele utilizarse para referirse a estas herramientas es CADE (*Computer Aided Document Engineering*), propuesto por Littleford (3). Según autores como Bieber y Vitali (4), el auge de Internet y la decisión de muchas organizaciones de adoptar la World Wide Web como infraestructura para sus aplicaciones, favorecerá en el futuro el desarrollo de los entornos CADE.

En la Universidad de Alcalá, el Grupo de Investigación en Ingeniería de la Información y de la Documentación, perteneciente al Departamento de Ciencias de la Computación, está desarrollando una herramienta de este tipo (figura 1), denominada ToolCADE (5). Su primera versión permite el diseño gráfico de la estructura y organización de los documentos de un sistema de información documental; también ofrece la posibilidad de modelar el aspecto dinámico de la documentación de dos formas diferentes: diseñando el flujo de trabajo a realizar en el seno de una organización, en el que estarán implicados los documentos; o modelando el ciclo de vida de cada documento, en forma de estados en los que puede encontrarse un documento, desde su creación hasta su destrucción.

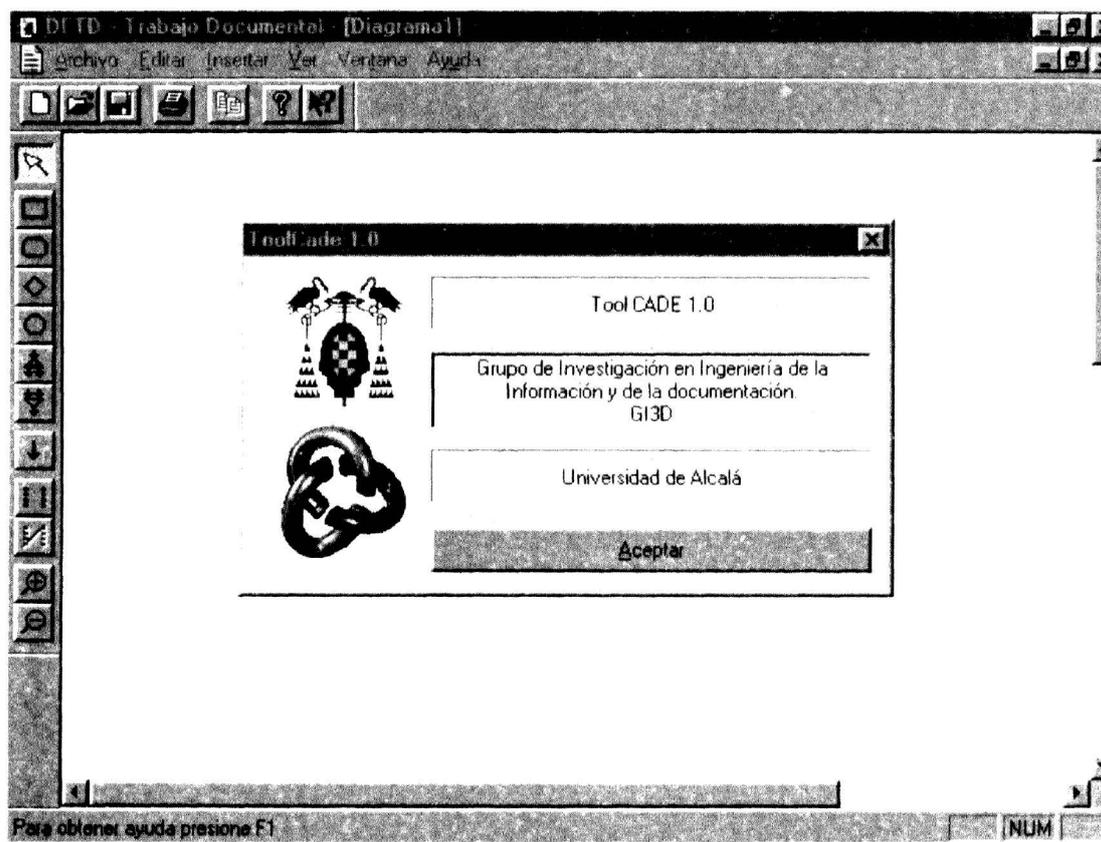
2 El origen: La ingeniería documental estructurada

Desde principios de los años setenta, en el ámbito de la producción de software se fue imponiendo gradualmente la evidencia de que el desarrollo de sistemas de infor-

mación complejos debía hacerse según un método riguroso de trabajo, de una manera sistemática y científica en lugar de la tradicional forma artesanal, la cual todavía sigue siendo el *modus operandi* de, cada vez menos, organizaciones de software. Actualmente, esta artesanía ha dado paso a lo que se conoce como Ingeniería de la Información o Ingeniería del Software, disciplina que considera el desarrollo de sistemas de información como un proyecto de ingeniería que implica un análisis riguroso de los requisitos que debe satisfacer el sistema para, posteriormente, proceder al diseño detallado del mismo y, finalmente, a su construcción e implantación en la organización usuaria (6).

A partir de la segunda mitad de la década de los ochenta y principios de los noventa, este enfoque empieza a ser asumido en el ámbito de la Documentación, ya que si, en general, la ingeniería implica un conjunto de métodos, técnicas y herramientas para el diseño y construcción de productos diversos, también el desarrollo de productos documentales de cierta complejidad, especialmente de tipo hipermedia (como enciclopedias temáticas o libros multimedia), debe abordarse utilizando un enfoque ingenieril, que suponga el establecimiento de un método riguroso de trabajo basado en una descomposición del proceso de desarrollo documental en fases y etapas, y en la

Figura 1
Herramienta de Ingeniería Documental ToolCADE.



aplicación de técnicas especialmente concebidas para cada una de las tareas de los proyectos de desarrollo de sistemas de información documental (7).

Como cualquier otra ingeniería, la documental también supone la ejecución de una serie de actividades para el desarrollo de sus productos, en este caso Sistemas de Información Documental. Estas actividades implican la planificación, el diseño, la producción y el mantenimiento de los documentos que constituyen el sistema, así como el análisis de su contenido, hacerlos más controlables y recuperables durante su explotación y mantenimiento (8).

Un proyecto de ingeniería documental comienza con el análisis de las necesidades de documentación de la organización usaria del sistema a desarrollar, elaborándose un plan en el que se establece el tipo de documentos a elaborar, la tecnología que se utilizará y la secuencia de operaciones para la producción y distribución de los documentos, y para la construcción de los lenguajes documentales que se utilizarán para la indización y recuperación de los documentos. Otra actividad fundamental en este tipo de proyectos, que le confiere su carácter de proyecto de ingeniería, es la que se refiere al diseño del sistema (2), debiéndose elaborar diferentes modelos conceptuales que representen la estructura estática y dinámica de los documentos, así como las características de su interfaz con el usuario (9).

Las primeras propuestas de metodologías de desarrollo de Sistemas de Información Documental (SID) como actividad de ingeniería se basaban en el denominado enfoque clásico o estructurado. Así, autores como Codina (10), De Moya (11) o Tramullas (12) proponen la aplicación de métodos de análisis estructurado, como el de Yourdon (13). Según este enfoque, antes de construir un SID ha de realizarse un modelo abstracto o conceptual del sistema en el que se representen los documentos y los procesos de transformación que tienen lugar en el interior del sistema.

En general, la principal característica de los métodos estructurados consiste en la separación que en ellos se hace de la información (documental) y de los procesos de transformación de ésta. Sin embargo, desde finales de los años ochenta existe una tendencia a considerar que ambos aspectos no pueden desligarse, sino que deben estar integrados, para obtener modelos más intuitivos del funcionamiento de los sistemas. Esta nueva forma de pensar es lo que se conoce como «orientación a objetos» (14). Este enfoque se basa en suponer que cualquier sistema, incluido el mundo real, se compone de objetos (personas, libros, muebles, vehículos, etc.), y que cada objeto posee intrínsecamente unos atributos y tiene asociadas unas determinadas funciones (por ejemplo, los atributos de una persona serían su nombre, edad, domicilio, etc., y las funciones comer, dormir, estudiar, etc.), de tal forma que cualquier proceso que ocurra en el sistema se puede concebir como una secuencia de interacciones entre todos los objetos que lo componen, realizando cada uno de ellos las funciones que tenga asignadas.

3 La concepción de un documento como objeto

La mayoría de los autores más representativos y cualificados en el ámbito del desarrollo de sistemas de información han ido abandonando el enfoque clásico estructurado para incorporarse a la corriente de la orientación a objetos. Éste ha sido, por ejemplo, el caso de alguien tan importante en este campo como Yourdon, quien en 1995 publicó una completa metodología de desarrollo de sistemas de información orientada

a objetos (15). Otro caso significativo en este sentido es el de las grandes compañías informáticas, como Microsoft, Hewlett-Packard, Oracle, MCI, Unysis o IBM, entre otras, que en 1997 definieron conjuntamente un método universal de modelado de sistemas de información, denominado UML (*Unified Modeling Language*), basado en el enfoque de la orientación a objetos (16).

También las administraciones públicas están modificando sus metodologías «oficiales» de desarrollo de sistemas de información para adaptarse a la orientación a objetos. Así ha ocurrido, por ejemplo, en Francia con la metodología MERISE, en el Reino Unido con SSADM, y en España con MÉTRICA, cuya nueva versión, como anuncia el Ministerio para las Administraciones Públicas, incluirá este enfoque (17).

En el contexto de los sistemas de información documental, también se está produciendo esta transición, especialmente en el caso de la documentación hipermedia. Ya se han establecido diferentes metodologías orientadas a objetos, como es el caso de las denominadas EORM (18) y OOHDM (19).

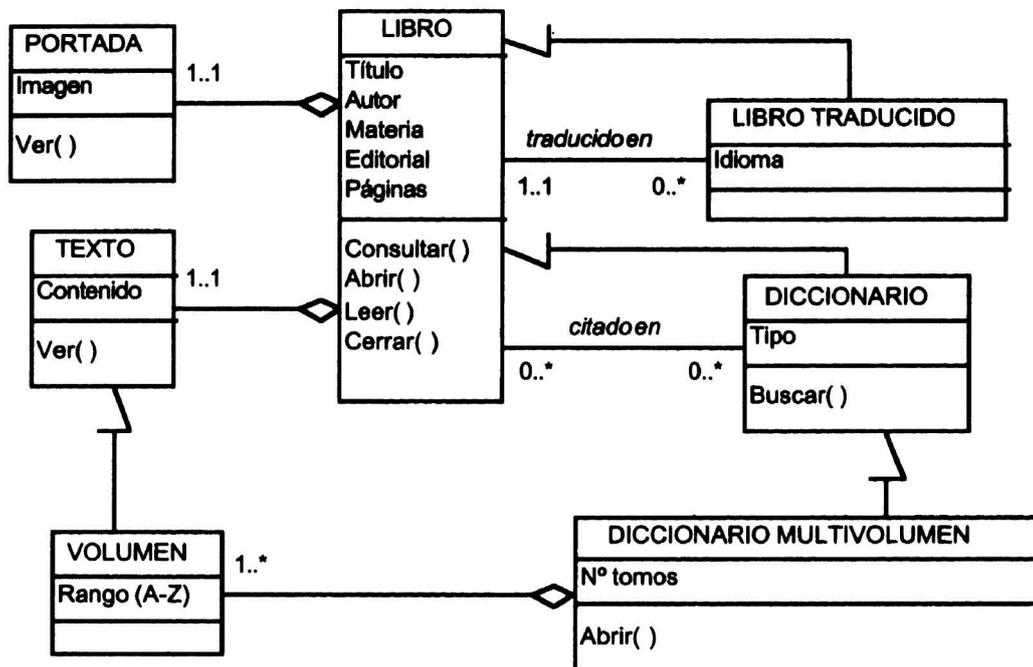
La orientación a objetos comienza también a ser considerada en las bases de datos documentales, en lugar del tradicional enfoque relacional. Así, productos comerciales de gestión de bases de documentos, como Lotus Notes, se basan en este enfoque, considerando los documentos almacenados como objetos que tienen asociadas operaciones que se pueden codificar en un lenguaje de programación orientada a objetos como LotusScript o Java.

Siguiendo el nuevo enfoque, el modelo conceptual de un sistema documental consistiría en una representación de las clases de objetos documentales implicados en el sistema, de sus propiedades o atributos, de las operaciones que se pueden realizar sobre ellos, y de las relaciones semánticas entre tipos de documentos. En la figura 2 puede observarse un modelo muy simplificado de un sistema correspondiente a una biblioteca digital, realizado con la técnica denominada *Diagrama de Clases*, definida en el estándar UML (16). En este contexto, se entiende por objeto documental un documento o estructura de documentos concretos (por ejemplo, un ejemplar de «El Quijote» o del «Diccionario de la Real Academia de la Lengua» compuesto de varios volúmenes), y por clase la representación de un grupo de objetos documentales con propiedades, estructura, funcionalidad y relaciones similares (por ejemplo, en la figura 2, la clase «Libro», a la que pertenecería «El Quijote», o la clase «Diccionario multivolumen» a la que pertenecería el de la Real Academia). Según esto, un objeto no será sino una instancia o ejemplar de una determinada clase.

En cuanto a los atributos y operaciones asociadas a una clase, los primeros representan propiedades de los objetos de esa clase, que asumirán un valor concreto para cada objeto en particular (por ejemplo, «Autor» podría ser un atributo de la clase «Libro», con un valor «Cervantes» para el caso concreto de «El Quijote»), mientras que las operaciones representan todo lo que se puede hacer con un objeto de esa clase (por ejemplo, «Abrir» un «Libro» o «Buscar» un término en un «Diccionario»).

Por otro lado, las relaciones entre clases hacen referencia a posibles vinculaciones entre determinados objetos de esas clases (por ejemplo, podría existir una relación llamada «citado en» entre la clase «Libro» y «Diccionario multivolumen», para indicar que algunos libros pueden citarse en algunos diccionarios, como sería el caso de una referencia a «El Quijote» desde el «Diccionario de la Real Academia de la Lengua»). Las relaciones entre objetos documentales también pueden utilizarse para representar que algunas clases de documentos son, en realidad, versiones refinadas de otras (por

Figura 2
Ejemplo de modelo de estructura documental



ejemplo, un «Diccionario multivolumen» podría ser un caso particular de «Diccionario»; o para indicar que un objeto se compone de otros objetos (en la figura 2, un «Libro» se compone de un objeto «Portada» y de un objeto «Texto»).

4 Interacciones entre objetos documentales

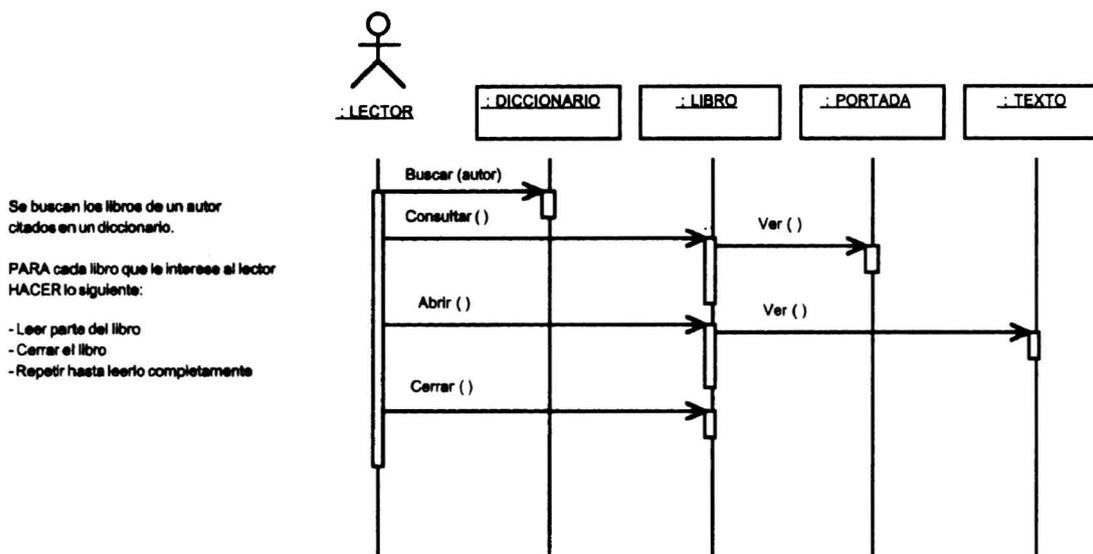
Aunque la estructura de un sistema de información documental se puede modelar mediante diagramas como el indicado en la figura 2, para tener una visión general de la funcionalidad del sistema conviene utilizar alguna otra técnica, también basada en el enfoque de la orientación a objetos, que muestre la secuencia de interacciones que se producen entre los objetos documentales cuando se lleva a cabo alguna de las funciones que ofrece el sistema.

En la figura 3 se representa, utilizando la técnica de UML denominada *Diagrama de Secuencia* (16), la interacción entre los objetos documentales que componen el sistema de la figura 2, cuando se trata de realizar la función de lectura del contenido de un libro citado en un diccionario. Para ello, en primer lugar, el usuario activaría la operación «Buscar Autor» del objeto «Diccionario»; cuando conoce el título del libro de ese autor que le interesa, activa la operación «Consultar» del objeto «Libro». Como consecuencia de la operación «Consultar», este objeto activa automáticamente la operación «Ver» de su objeto asociado «Portada», para que el lector pueda contemplarla y decida si le interesa abrir el «Libro»; si es así, activa la operación «Abrir», que desencadena, a su vez, la activación de la operación «Ver» para visualizar el texto del li-

bro. Finalmente, cuando el usuario finaliza la lectura, activaría la operación «Cerrar» del objeto «Libro».

Lo que se representa en diagramas como el de la figura 3 es lo que ocurre en un sistema de la forma más intuitiva posible, ya que es así como funciona cualquier sistema del mundo real, que está formado por objetos que participan juntos en la consecución de objetivos comunes: las funciones del sistema.

Figura 3
Diagrama de interacción documental correspondiente a la lectura de un libro citado en un diccionario.



5 Conclusiones

Desde que Otlet sugiriese la posibilidad de que en el futuro se automatizase el proceso documental han transcurrido más de seis décadas (20). En este tiempo se ha consolidado la Documentación como disciplina científica, pero todavía no como actividad susceptible de ser llevada a cabo según un proceso de ingeniería. La última década del siglo XX se está caracterizando en este ámbito por un cambio de actitud por parte de los productores de sistemas de información documental, que ya empiezan a incorporar herramientas de diseño que permiten realizar modelos conceptuales de los sistemas a desarrollar, para comprender mejor lo que se está construyendo y facilitar en el futuro la labor de mantenimiento de esos sistemas (1, 2).

Igual que ha ocurrido en el campo de la Ingeniería de la Información, la orientación a objetos comienza a imponerse como el enfoque más adecuado para el diseño y construcción de los sistemas de información documental, siendo ya una realidad las metodologías documentales basadas en este enfoque. En este sentido, gracias a Internet se está avanzando en este campo, ya que la mayor parte de los métodos de ingeniería documental orientada a objetos surgen precisamente para el desarrollo de sistemas de información documental en la red, debido a que la World Wide Web está

suponiendo un cambio en la forma de concebir la producción de aplicaciones informáticas para Internet (21), que está pasando de ser considerado un proceso de construcción de software, a convertirse en un proceso de ingeniería documental, teniendo en cuenta que lo que en realidad se produce son documentos en forma de páginas Web, aunque con software incorporado.

Para mantenerse informado sobre los trabajos que se están llevando a cabo en el mundo sobre estos temas, se sugiere la búsqueda en la Web de información sobre «Ingeniería Documental», «Document Engineering», «Hypermedia Design» y «CADE».

6 Bibliografía

1. BRERETON, P.; BUDGEN, D., y HAMILTON, G. Hypertext: The Next Maintenance Mountain, *Computer* (New York), 1998, diciembre, p. 49-55.
2. SUTTON, M. J. *Document Management for the Enterprise: Principles, Techniques and Applications*. New York: John Wiley & Sons, 1996.
3. LITTLEFORD, A. Artificial Intelligence and Hypermedia. En: Berk, E.; Devlin, J., editores. *Hypertext/Hypermedia Handbook*. New York: McGraw-Hill, 1991, p. 365-378.
4. BIEBER, M. y VITALI, F. Toward Support for Hypermedia on the World Wide Web. *Computer* (New York) 1997, junio, p. 62-70.
5. LÓPEZ, L.; GONZÁLEZ, L. A.; HILERA, J. R. y MARTÍNEZ, J. M. ToolCADE: A Computer Aided Document Engineering Environment. *DLM-Forum on Electronic Records*, 1999, octubre, Bruselas, 18-19.
6. MARTÍNEZ, J. M.; LÓPEZ, L.; CASTAÑO, B.; MALPICA, J. A.; HILERA, J. R. y GUTIÉRREZ, J. A. *Metodología de Desarrollo de Sistemas de Información*. Alcalá de Henares: Universidad de Alcalá, 1995.
7. GARZOTTO, F.; MAINETTI, L., y PAOLINI, P. Hypermedia Design: Analysis and Evaluation Issues. *Communications of the ACM* (New York) 1995, vol. 38, n.º 8, p. 74-86.
8. SIMINIANI, M. *Intranets, empresa y gestión documental*. Madrid: McGraw-Hill, 1997.
9. DÍAZ, P.; CATENAZZI, N., y AEDO, I. *De la Multimedia a la Hipermedia*. Madrid: RAMA, 1996.
10. CODINA, L. Modelo conceptual de un Sistema de Información Documental. *Revista Española de Documentación Científica*, 1994, vol. 17, n.º 4, p. 440-449.
11. DE MOYA, F. *Los Sistemas Integrados de Gestión Bibliotecaria*. Madrid: ANABAD, 1995.
12. TRAMULLAS, J. Los Sistemas de Información: Una reflexión sobre Información, Sistema y Documentación. *Revista General de Información y Documentación*, 1997, vol. 7, n.º 1, p. 207-229.
13. YOURDON, E. *Modern Structured Analysis*. Englewood Cliffs (EE.UU.): Prentice Hall, 1989.
14. BOUZEGHOUB, M.; GARDARIN, G., y VALDURIEZ, P. *Objetos*. Barcelona: Ediciones Gestión 2000, 1996.
15. YOURDON, E.; WHITEHEAD, K.; THOMANN, J.; OPPEL, K., y NEVERMANN, P. *Mainstream Objects: An Analysis and Design Approach for Business*. Upper Saddle River (EE.UU.): Prentice Hall, 1995.
16. *Unified Modeling Language Documentation*. Santa Clara (EE.UU.): UML Resource Center, 1999. <http://www.rational.com/uml/resources/documentation>.
17. *Proyecto MÉTRICA Versión 3*. Madrid: Ministerio para las Administraciones Públicas, 1999. <http://www.map.es/csi/pg5m42.htm>.
18. LANGE, D. B. *An Object-Oriented Design Approach for Developing Hypermedia Information Systems (Research Report RT0112)*. Tokyo: IBM Tokyo Research Laboratory, 1995.

19. SCHWABE, D. y ROSSI, G. The Object-Oriented Hypermedia Design Model. *Communications of the ACM* (New York), 1995, vol. 38, n.º 8, p. 45-46.
20. OTLET, P. *Traité de Documentation. Le livre sur le livre. Théorie et pratique*. Bruselas: Editions Mundaneum, 1934.
21. CERI, S.; FRATERNALI, P., y PARABOSCHI, S. Design Principles for Data-Intensive WEB Applications. *Terceras Jornadas de Ingeniería del Software*, 1998, noviembre, 11-13, Murcia.

José Ramón Hilera González

Departamento de Ciencias de la Computación.

Facultad de Ciencias de la Documentación, Universidad de Alcalá,

28871 Alcalá de Henares (Madrid). Correo-e: ccjrhg@cc.alcala.es.

X ASAMBLEA GENERAL DE LA RED EUROPEA DE INFORMACION Y DOCUMENTACION SOBRE AMERICA LATINA (REDIAL)

Los días 24 a 26 de mayo de 2000 ha tenido lugar en la Universidad Internacional de Andalucía, Sede Iberoamericana de La Rábida, la X Asamblea General de REDIAL, con la asistencia de 25 personas representando a unos 30 centros de documentación y bibliotecas especializadas en temas latinoamericanistas de cuatro países: España, Francia, Suecia y Reino Unido. Por parte de España estaban representados la Biblioteca Nacional, la Biblioteca Hispánica, la biblioteca de la Escuela de Estudios Hispano Americanos del CSIC, el Centro de Estudios sobre Relaciones Internacionales de la Fundación Ortega, el Centro de Información Documental de Archivos (CIDA), IEPALA, el Instituto de Estudios de Iberoamérica y Portugal, el Instituto Internacional de la Gobernabilidad de Barcelona y el CINDOC. Del Reino Unido Canning House, representando también al ILAS y a la Hispanic Section de la British Library; De Suecia, el LAI de Estocolmo, y de Francia, estuvo representado el Reseau Amérique Latine y varias de las instituciones que lo forman: IHEAL, GRAL, BDIC, CEDOCAL, además de la Biblioteca del CADIST de la Universidad de Burdeos.

Tras un breve repaso de la situación de los estudios americanistas en cada uno de los países y de las actividades más notables de cada uno de los centros representados, se pasó a hacer un balance de la situación de los trabajos desarrollados por los grupos. Seleccionamos las actividades de mayor interés:

El grupo de *Salvaguardia de archivos de organizaciones populares* está aún ocupándose de microfilmear los archivos del MNR, en Bolivia. Hay el proyecto de microfilmear los archivos del APRA en Perú y los archivos del sindicato de mineros de Potosí en Bolivia con documentos entre 1930 y 2000. Como siempre, al finalizar los proyectos se entrega una copia al Archivo de la Nación y a la institución cuyos archivos se microfilman que, por supuesto, conserva también los originales. Las instituciones participantes se quedan también con una copia que entregan al Archivo Histórico

o Nacional de su país. Participan BDIC y CESAME de Francia, CIDA de España, la Fundación Feltrinelli de Italia y el Instituto de Historia Social de Amsterdam

El grupo de trabajo sobre *Etnias* informó sobre la evolución de los trabajos desde su inicio en 1998: se están realizando dos bases de datos en paralelo: una en Francia, ETHNIES, que cuenta con 1.000 entradas admitidas y 400 variantes de diferentes etnias, con descripción histórico-geográfica de cada una y equivalencias en los Subject Headings de la Library of Congress, y la base de datos española ETNIAS, que tiene 500 entradas y que de momento trata de recoger la terminología empleada en la indexación de la literatura específica por cada una de las siguientes instituciones: BNE, bibliotecas del CSIC, IEPALA, Biblioteca Hispánica y CINDOC, incluyendo los términos elegidos, los equivalentes y los relacionados. Se aporta también la denominación utilizada en LCSH. A lo largo del próximo año, se intentará crear un producto que ofrezca a los investigadores y antropólogos especialistas en Antropología americana, las distintas denominaciones de las etnias y sus variaciones. Participan en este trabajo la C.BIC, IEPALA, BNE, Biblioteca Hispánica, IHEAL de Francia y CINDOC.

El grupo de trabajo de *Sumarios de revistas latinoamericanistas europeas* ha trabajado para conseguir que cada país aporte de manera sistemática la digitalización de la portada de las revistas latinoamericanistas más importantes y los sumarios. Por el momento hay incorporadas revistas francesas y españolas, cuyos sumarios estarán en breve disponibles en <http://red-redial.org> y que de momento residen en el sitio web de la red francesa <http://www.reseau-amerique-latine.fr/index2.html>. Esta base de datos de sumarios se actualiza cuatro veces al año. Se proyecta añadir informaciones sobre localización de las revistas en los centros de la red para así facilitar el acceso a los textos.

Creación de un índice de recursos de información de y sobre América Latina en Internet: Todos los participantes han estado de acuerdo en que este proyecto es una apuesta estratégica para REDIAL. Se pretende que sea realmente un punto de referencia inexcusable para la investigación latinoamericanista. Aun está en sus inicios, pero ya cuenta con 500 enlaces, aunque únicamente están incorporados los recursos disponibles en las sedes virtuales de España, miembros de REDIAL, parte del resto de Europa y de Estados Unidos, además de haberse indizado los recursos de Chile, Colombia, Argentina, Perú, Bolivia, Costa Rica y Puerto Rico. En su etapa inicial se ha partido de la recopilación de Graciela Fainstein Lamuedra, publicada por el CINDOC en 1998 y se ha contado con la colaboración técnica del equipo de Cibermetría del CINDOC. El trabajo de coordinación lo ha llevado a cabo el CERI que ha contado para ello con una ayuda del Ministerio de Educación y Cultura. El proyecto es ambicioso y será la ventana por la que REDIAL se asomará al mundo de Internet. Lo ya realizado puede consultarse en <http://www.red-redial.org>.

Además de estos proyectos, siguen desarrollándose otros de interés como la base de datos de Tesis Europeas sobre América Latina o los Directorios de Americanistas que siguen actualizándose en diferentes países. A ello hay que añadir en el momento actual la preparación de los trabajos para el próximo Congreso Internacional de Americanistas, a celebrar en Varsovia en el próximo mes de julio: REDIAL organiza un simposio en el que se pondrán en común las diferentes perspectivas de los *estudios sobre la ciudad latinoamericana* que se realizan en Europa y en el continente americano.

Quizás la conclusión general de la Asamblea es la convicción de que REDIAL tiene un espacio dentro del panorama de la investigación y la información sobre América Latina en Europa y que la utilización eficaz y cooperativa de los recursos y en especial de la sede virtual, puede dar conexión y visibilidad al trabajo desarrollado.

Adelaida Román
CINDOC-CSIC