

# TDC@t

## Acceso electrónico a las tesis doctorales de Cataluña

[Nota: aquest article va ser publicat dins [El Profesional de la informació](#), vol. 11, nº1, enero-febrero 2002, pp. 28-34.  
Pàgina web de la revista a <http://www.szp.swets.nl/szp/journals/pi.htm>]

Lluís Anglada\*, Ingrid Bárcena\*\*, Joan Cambras\*\*, Núria Comellas\*, Miquel Huguet\*\*, Ramon Ros\*

\*Consorti de Biblioteques Universitàries de Catalunya (CBUC)

\*\*Centre de Supercomputació de Catalunya (CESCA)

### La difusión de las tesis doctorales

Las tesis doctorales son uno de los resultados más visibles de las tareas de investigación que toda universidad lleva a cabo. Al mismo tiempo son frecuentemente el punto de partida de trabajos posteriores, iniciando así un proceso en cadena que refleja y asegura la vida intelectual de una sociedad determinada. Tradicionalmente, el tratamiento documental que se ha dado a las tesis doctorales ha sido el que suelen recibir los documentos que conocemos bajo la denominación genérica "literatura gris", la cual engloba a los trabajos de alto interés para la comunidad investigadora pero de baja visibilidad a causa de que no acostumbra a seguir los circuitos editoriales normales y de que no existen demasiadas copias impresas en circulación.

Estas dificultades de acceso a las tesis doctorales representan un obstáculo para la difusión de los avances científicos de un país. Hasta ahora la edición en microficha de las tesis ha paliado en cierto modo este problema. Esta técnica aporta mejoras en la conservación y difusión respecto a la copia en papel, pero podemos decir que actualmente ha sido superada por el hecho de que la gran mayoría de las tesis han sido generadas originalmente en soportes digitales mediante programas de autoedición, procesadores de textos, etc. Esta situación abre nuevas posibilidades de tratamiento documental de las tesis doctorales y facilita en gran manera su difusión por las autopistas de la información.

A nivel internacional, se han desarrollado diversas iniciativas para crear servidores de tesis digitales en universidades europeas y norteamericanas, aunque en casi todos los casos se trata de iniciativas individuales que no se

han difundido adecuadamente. En Europa, han surgido algunos intentos en Alemania, Holanda y Austria, pero curiosamente no se ha aprovechado la existencia de diversos proyectos de colaboración en bibliotecas digitales (como el Telematics for Libraries) para poner en marcha un programa colectivo de digitalización de tesis. De todas maneras cabe destacar que Austria es quizá el único país europeo que ha dedicado esfuerzos en crear un programa a nivel nacional para digitalizar todas las tesis que se presenten en sus universidades.

En España, la situación hasta hace poco ha sido muy parecida a la del resto de Europa. Tan sólo cabe remarcar las iniciativas por parte de algunas universidades (Complutense, Lleida, León) de editar y distribuir algunas tesis en CD-ROM. En la Universidad Politécnica de Valencia se han digitalizado algunos proyectos de final de carrera que sólo son accesibles desde la misma universidad.

En los Estados Unidos de América, además de proyectos individuales llevados a cabo por diversas universidades y de productos comerciales como la base de datos "Dissertation Abstracts" de UMI, cabe destacar la iniciativa de mayor alcance en el ámbito de las tesis doctorales electrónicas: *Networked Digital Library for Theses and Dissertations (NDLTD)*, proyecto iniciado a principios de los noventa por la Virginia Tech University en los EUA y al cual se están añadiendo muchas universidades de otros países. NDLTD ha desarrollado procedimientos y normativas (así como un software) que están sentando precedente mundial en diversos aspectos del tratamiento de las tesis doctorales digitales. Hasta el momento, 120 universidades de todo el mundo se han asociado a NDLTD<sup>1</sup>.

## **El modelo NDLTD**

Las tesis en NDLTD son accesibles mediante un sistema que incorpora las características de cualquier sistema bibliográfico basado en un motor de búsqueda. El usuario puede recuperar la tesis por autor, título o departamento donde se ha presentado. A continuación aparece un registro que a estas informaciones añade un resumen, una URN (Uniform Resource Name: nombre lógico que enlaza con una dirección física) y el nombre del director de la tesis, entre otros datos. La mayoría de las tesis dispone de un fichero con el texto completo de la misma, que se puede por lo tanto leer sin restricciones. En el caso de Virginia Tech (universidad impulsora de NDLTD) las tesis están almacenadas en un servidor centralizado.

A medio plazo se proyecta implementar un sistema basado en "búsqueda federada". Esto significa que una misma búsqueda se ejecutaría en todas las bases de datos de la red NDLTD; el resultado de la búsqueda sería un listado de las tesis que cumplieran las premisas de la búsqueda, sin importar el servidor (universidad) donde estuvieran localizadas.

---

<sup>1</sup> Para datos más extensos, vid. D-lib magazine, vol. 7 (sept 2001) n.9, <http://www.dlib.org/dlib/september01/09contents.html> (consulta: 19.09.01)

De todos modos, los aspectos técnicos relativos al sistema informático son sólo uno de los aspectos a tener en cuenta. Otro no menos importante, es el de la creación de los circuitos internos que posibiliten el flujo de comunicación entre el doctorando y la universidad, con el objetivo final de dar acceso público al texto completo de la tesis. En la Virginia Tech, es el propio doctorando quien publica su tesis en NDLTD. Para ello, la universidad programa habitualmente cursos para los estudiantes de doctorado orientados a la redacción de tesis y a su preparación para la publicación electrónica de las mismas. Según este modelo, el estudiante es responsable de entregar todos los datos y ficheros relativos a su tesis; en consecuencia, esto reduce el papel de la biblioteca/centro de informática a la simple comprobación de que los datos son correctos y verdaderos, a describir la tesis en el catálogo de la biblioteca, y a darle acceso electrónico desde el servidor.

Uno de los temas más delicados con respecto a la publicación de las tesis electrónicas en Internet es el de los derechos de autor. Existen diversas soluciones que intentan resolver este tema, armonizando en todo momento el derecho del autor a proteger su obra con el derecho de cualquier ciudadano a acceder a la información contenida en la tesis. La tendencia generalizada es la permitir el acceso universal al texto completo de las tesis sin restricción alguna, aunque en algunos casos especiales (tesis que luego se reconvierten en patentes, tesis que incluyen textos para los que el autor no posee los derechos de reproducción, etc.) existen soluciones alternativas: limitar el acceso sólo a la red de la universidad donde se presentó la tesis, o bien permitir la visualización sólo de los datos del registro (autor, título, etc.) y del resumen de la tesis. En cualquier caso, el autor y la universidad siempre suscriben un contrato legal donde se especifican los usos autorizados de la tesis, así como los derechos y deberes de ambas partes.

### **Las tesis doctorales electrónicas en Cataluña**

En 1999 el entonces Comisionado para la Sociedad de la Información de la Generalitat de Catalunya<sup>2</sup> y LOCALRET<sup>3</sup> promovieron un Plan estratégico para impulsar y desarrollar la sociedad de la información en Cataluña<sup>4</sup>. El Plan se realizó con la participación de un amplio espectro de expertos y hizo una serie de recomendaciones entre las cuales algunas tenían por agentes las universidades. Para activar la realización de las prioridades establecidas en el Plan, el Departamento de Universidades, Investigación y Sociedad de la Información de la Generalitat promovió la firma de un convenio por el que el propio Departamento, las universidades el CESCA y el CBUC se comprometían a realizar ciertas acciones a favor del fomento de las nuevas tecnologías<sup>5</sup>.

---

<sup>2</sup> Actualmente es una Secretaría dentro del Departamento de Universidades, Investigación y Sociedad de la Información

<sup>3</sup> LOCALRET es un consorcio que agrupa la mayoría de los municipios de Cataluña y que tiene por finalidad promover el desarrollo y uso de nuevas tecnologías.

<sup>4</sup> Catalunya en xarxa: Pla estratègic per a la societat de la informació (Barcelona: Comissionat per a la Societat de la Informació, 1999).

<sup>5</sup> Teraflop, nº44 (oct 99), pp. 6-7.

Una de ellas era el diseño de un servidor de tesis doctorales en formato digital (TDC@t) para las universidades de Cataluña. Mediante la consulta remota se pretendía incentivar el uso de las tesis y facilitar su conservación.

En este contexto, las universidades de Barcelona, Autònoma de Barcelona, Politècnica de Catalunya, Pompeu Fabra, Girona, Lleida, Rovira i Virgili y Oberta de Catalunya, como entes generadores de unas mil nuevas tesis doctorales cada año, delegaron en el CBUC la coordinación de un proyecto de edición, almacenamiento y acceso electrónicos de las tesis catalanas. El CESCO participa en el proyecto en los aspectos relativos a la implementación y funcionamiento del servidor que gestiona la base de datos TDC@t. La Biblioteca de Catalunya participa reuniendo en su sede, mediante el depósito legal, una copia de archivo de las tesis doctorales publicadas en TDC@t. De esta manera se asegura la conservación y preservación futura de estos documentos. El programa de acceso a tesis se inscribe en la Biblioteca Digital de Catalunya, la cual tiene por objetivo el ofrecer un núcleo básico de información electrónica accesible a toda la comunidad universitaria e investigadora de Cataluña.

En 1999, el CBUC encargó un informe<sup>6</sup> sobre la situación de la digitalización de las tesis doctorales en el mundo. Fruto de este primer estudio, se decidió iniciar las tareas de organización entre las entidades implicadas, así como las consultas legales para todo lo referente a la edición y distribución en línea de las tesis. Así mismo, a principios de diciembre de 2000 se acordó la adhesión de TDC@t al proyecto NDLTD.

## **El proyecto TDC@t**

Los objetivos de TDC@t están inspirados en los logros de NDLTD y se formularon de la siguiente manera:

- disponer de un software para acceder y visualizar las tesis doctorales en formato electrónico, a texto completo
- generar un conjunto de procedimientos y directrices sobre cómo redactar tesis electrónicas para su posterior publicación en Internet
- ofrecer asesoramiento a los estudiantes para publicar sus tesis en formato electrónico según los requisitos del proyecto
- armonizar el acceso público a las tesis doctorales electrónicas con las leyes de propiedad intelectual y los derechos de los autores
- lograr un compromiso de los gestores universitarios y crear una plataforma de colaboración entre docentes, personal de servicios, estudiantes y bibliotecas para garantizar el acceso y la conservación a las tesis doctorales electrónicas.

---

<sup>6</sup> DIJKSTRA, J.M.M. Towards storage and access of electronic theses: A proposal of its organisation for the Consortium of Catalan Libraries. Barcelona: [s.n.], 1999 [Manuscrito]

Pese a estar inspirado en gran medida en NDLTD, TDC@t difiere del modelo en un aspecto organizativo básico: en TDC@t la universidad es la encargada de introducir la información descriptiva y el fichero de la tesis en la base de datos. Ello implica que los canales de comunicación, tanto externos como internos, de la universidad deben estar bien definidos de antemano para que, en el momento en que el autor apruebe su tesis doctoral, ésta entre en el circuito de digitalización de tesis con total naturalidad y a la mayor brevedad posible. Cada universidad participante en TDC@t tiene total libertad para establecer este circuito interno del modo que le sea más operativo, si bien todas ellas deben asegurar que las tesis que entraran a formar parte de TDC@t cumplen unos requisitos mínimos comunes. Estos requisitos son de diversa índole: legal, organizativo e informático.

Los requisitos legales son básicamente dos: la universidad firma un convenio de colaboración con el CBUC y el CESCA mediante el cual se determinan las relaciones de ambas instituciones en lo referente a TDC@t, y el autor firma con la universidad un contrato de edición que establece los derechos y deberes de ambas partes así como los usos autorizados de la tesis doctoral. El CBUC contó con el asesoramiento de un bufete de abogados especializado en edición electrónica para redactar ambos documentos, dado que este es un tema nuevo sobre el cual la legislación vigente no ofrece una casuística clara.

Los requisitos organizativos implican que cada universidad debe designar un interlocutor responsable del funcionamiento de TDC@t en su institución. Este interlocutor es la persona de contacto con el CBUC, desde donde se le facilita cualquier información o documento que le pueda ser útil en alguna fase de desarrollo del proyecto: difusión, formatos informáticos aceptados, entrada de datos en TDC@t, dudas de procedimiento, etc. Dado que la biblioteca de la universidad deberá catalogar la tesis electrónica, también es necesario designar un interlocutor en la biblioteca que se responsabilice de describir en el catálogo toda nueva tesis que entre en el sistema.

### **Soporte informático**

El proyecto TDC@t dispone de un servidor central desde donde se almacenan y difunden las tesis y que sirve también para sustentar el proceso administrativo de la introducción de los documentos, de forma distribuida, por parte de cada universidad. Esta máquina corresponde actualmente a un HP Exemplar V2500 con 16 procesadores PA8500. El servidor es uno de los supercomputadores de que dispone el CESCA y, además de TDC@t, está dedicado al cálculo científico.

El programa que gestiona TDC@t es una adaptación del paquete estándar cedido en el marco de NDLTD y desarrollado por la universidad Virginia Tech. El mayor cambio ha sido la adaptación a un proyecto formado por ocho universidades (el software original estaba diseñado para una sola) conservando la independencia y seguridad de datos de cada una pero mostrando al usuario una sola base de datos compacta y buscable en su conjunto. Otros cambios

que han requerido un esfuerzo importante, han sido la sustitución del buscador por otro más genérico y potente y el cambio parcial en la estructura de datos.

Respecto a la estructura y formato de los datos, se ha elegido el formato PDF. Este formato, aunque propietario de Adobe Inc., es ya un estándar de facto en la publicación electrónica por Internet con las ventajas de ofrecer cierta protección (una tesis depositada en TDC@t no puede modificarse directamente) y conserva la presentación y formato de la tesis original con independencia del dispositivo utilizado por el usuario para su consulta.

En una segunda fase, este software permitirá la generación automática de algunos metadatos (además de los ya existentes) que darán lugar a la posibilidad de la búsqueda federada dentro del conjunto de NDLTD. Al mismo tiempo las tesis son catalogadas en el Catàleg Col·lectiu de les Universitats de Catalunya (CCUC) y desde allí copiadas en los catálogos locales respectivos. La consulta de un registro bibliográfico de una tesis incluida en TDC@t permite enlazar con el texto completo de la tesis.

## **Beneficios y valoraciones finales**

Aunque aún es pronto para hacer una evaluación del incremento de difusión de las tesis disponibles en TDC@t, los datos disponibles a nivel internacional<sup>7</sup> muestran que las tesis accesibles en línea son consultadas con una intensidad mucho mayor que lo que lo eran en formato papel o microforma. Los beneficios de este hecho para la comunidad universitaria e investigadora son innegables. El lector puede encontrar datos relevantes de una especialización muy alta que de otra forma quedarían inaccesibles. El autor gana a su vez lectores, rentabilizando así su esfuerzo de investigación. Finalmente, la investigación realizada en el marco de los programas de doctorado adquiere una relevancia que anteriormente no tenía.

A octubre del 2001, casi dos años después de formular estos objetivos y seis meses después de la carga de la primera tesis, TDC@t está en marcha y dispone de una base de datos (<http://www.tdcat.cbuc.es>) con más de doscientas tesis doctorales a texto completo. Para llegar hasta aquí ha sido necesario tomar decisiones tanto en el aspecto técnico como en el organizativo, trabajando a diversos niveles (gestores universitarios, técnicos informáticos, jefes de servicios de publicaciones, bibliotecarios, etc.).

Al hablar de edición electrónica se suele poner el acento en los temas técnicos. En el caso de TDC@t estos han representado los menores problemas. Los principales se han encontrado en la tarea de cambiar circuitos de procesos ya establecidos (el de la edición de tesis en microforma) y, especialmente, en la coordinación de grupos de personas que normalmente no trabajan juntos. La nueva forma de difusión de las tesis ha supuesto en muchos casos cambios en la normativa de las comisiones de doctorado y de tercer ciclo así como la preparación de contratos para la edición electrónica. La preparación y

---

<sup>7</sup> vid. D-lib magazine, vol. 7 (sept 2001) n.9, <http://www.dlib.org/dlib/september01/09contents.html> (consulta: 19.09.01)

funcionamiento del proyecto ha supuesto llegar a acuerdos entre los servicios académicos, de ediciones y las bibliotecas de ocho universidades así como tareas de modificación de los circuitos de entrega del depósito legal por parte de la Biblioteca de Catalunya.

Por otra parte, este aparente punto débil, que ha frenado los plazos de implementación, ha resultado ser el mejor punto de apoyo para conseguir resultados a nivel colectivo. Ciertamente, los obstáculos encontrados en diferentes puntos han podido ser superados gracias a que alguien anteriormente había encontrado una solución práctica para el problema. Uno de los elementos correctores introducidos en la planificación del proyecto fue la creación de una Intranet desde la que se pueden consultar los diferentes materiales de trabajo elaborados por cada universidad. Aquí se encuentran los contratos de edición, folletos de divulgación, formularios muestra de petición de ISBN, etc. La naturaleza colectiva del proyecto no permitió en su momento acordar un sistema común de indización o de clasificación de las tesis. Estas, a pesar de poder ser buscadas por diferentes conceptos, no pueden ser 'hojeadas' de una forma global. El número actual de tesis en el servidor (a pesar de ser aún muy inferior del esperado a corto plazo) ha evidenciado la necesidad de subsanar esta deficiencia.

Debe destacarse también, y esto ha estado siempre presente en el proyecto, que el trabajo conjunto permite una visibilización mucho mayor de las tesis de cada universidad. Las dimensiones diferenciadas de las universidades implicadas supone a su vez una producción desigual de tesis. A pesar de ello, poder acceder a ellas de forma conjunta ha permitido que la consulta del servidor sea más coste-efectiva y, por tanto, incrementar su uso.

Finalmente, proyectos como el ND LTD o TDC@t deben contemplarse como aspectos de dos movimientos generales asociados al fenómeno de las bibliotecas digitales y de la edición electrónica. En primer lugar nos referimos al amplio movimiento en pro de la recuperación de contenidos por parte de las universidades y centros científicos. Recordemos que los desarrollos de la comunicación científica han conducido a una situación en la que el autor, que suele situarse en una universidad, cede sus trabajos a editoriales científicas que a su vez los venden a las universidades a través de las bibliotecas. La recuperación de contenidos no tiene muestras solo en las tesis sino que presenta importantes realizaciones en servidores de 'preprints' y de revistas electrónicas. En segundo lugar, nos referimos al progresivo (pero más lento que lo vaticinado por algunos tecnólogos) paso de la edición en soportes tradicionales a la edición electrónica. Este proceso tiene diferentes velocidades y nichos pero es evidente que encuentra en el campo de los trabajos de investigación científica un terreno abonado para su desarrollo.

### **Más información:**

- TDC@t: <http://www.tdcat.cbuc.es>
- NDLTD: <http://www.ndltd.org>
- CBUC: <http://www.cbuc.es>
- CESCO: <http://www.cesca.es>
  
- Fox, Edward A. et al. "Networked Digital Library of Theses and Dissertations An International Effort Unlocking University Resources" [en línea]. En: D-Lib Magazine, Sep.1997 <<http://mirrored.ukoln.ac.uk/lis-journals/dlib/dlib/dlib/september97/theses/09fox.html>> [Consulta 27 junio 2001]
  
- McMillan, G. "Managing electronic theses and dissertations". En: College and Research Libraries News, May 2000, p. 413-414.