

# **INFORMÁTICA, DERECHO Y DOCUMENTACIÓN. EXPERIENCIAS Y POSIBILIDADES DE APLICACIÓN DE LOS LENGUAJES DE MARCADO DE TEXTO (SGML, HTML Y XML) A LOS DOCUMENTOS JURÍDICOS**

J. Tomás Nogales Flores, Bonifacio Martín Galán y M<sup>a</sup> del Carmen Arellano Pardo  
Dpto. Biblioteconomía y Documentación  
Universidad Carlos III de Madrid

## **PALABRAS CLAVE:**

Documentación Jurídica / legislación / Jurisprudencia / Lenguajes de marcado / SGML / HTML / XML / Internet / World Wide Web

## **RESUMEN:**

Los lenguajes de marcado de texto pueden ser contemplados como una tecnología complementaria o alternativa a las bases de datos documentales convencionales, presentando ventajas notables sobre éstas para la gestión de colecciones de documentos extensos, con estructura compleja y variable y abundantes referencias cruzadas, como es el caso de los documentos jurídicos. Se exponen experiencias de tratamiento y difusión de documentación jurídica en diversos países, haciendo uso de lenguajes de marcado como SGML, HTML y XML, así como las llevadas a cabo desde hace varios años por un equipo de trabajo del Departamento de Biblioteconomía y Documentación de la Universidad Carlos III de Madrid, que han tomado la forma de contratos de asesoría con empresas, proyectos de investigación con financiación pública o tesis doctorales.

---

## **Introducción**

Las interrelaciones que se han producido entre las ciencias del Derecho, la Informática y la Documentación a lo largo de estas últimas décadas son ya sobradamente conocidas. No pretendemos en esta comunicación, por tanto, incidir directamente sobre este aspecto sino, más bien, señalar algunas ideas esenciales de esta estrecha relación desde la aparición de los principales lenguajes de marcado de textos de aplicación en el espacio electrónico de la red Internet, donde la existencia de numerosas experiencias prácticas en todo el mundo demuestran la validez de estas tecnologías para el tratamiento y difusión de la información de carácter jurídico. En muchas de estas experiencias ha sido de vital importancia la incorporación de bibliotecarios y documentalistas constituyendo, por tanto, ese tercer pilar en el que se ha de sustentar el desarrollo de cualquier proyecto de gestión de la información jurídica, como se irá exponiendo a continuación.

Desde el punto de vista de los profesionales del Derecho en su relación con las otras ciencias aquí anunciadas, ha sido la Informática la que de forma más extensa y profunda ha sido abordada en la bibliografía científica.

La aplicación de la Informática dentro del ámbito jurídico se inicia poco tiempo después de que Norbert Wiener elaborase sus postulados sobre la cibernética, ya indicándose en la

propia obra de este autor la posibilidad de aplicar sus conclusiones al Derecho<sup>1</sup>. Pero será un juez norteamericano, Lee Loevinger, responsable de la División Antitrust, quien emplee por primera vez el término *jurimetría* en un artículo aparecido en 1949 en el número 33 de la revista *Minnesota Law Review*<sup>2</sup>. Este jurista establecía que la Jurimetría suponía el conjunto de investigaciones, tanto lógico-matemáticas como estadísticas, dirigidas a los distintos tipos de análisis de la información jurídica y a su tratamiento mecánico mediante computadoras, con la finalidad de documentar de forma automática dicha información, racionalizarla, enseñarla, administrarla y preservarla dentro de la esfera del Derecho.

Pero, sin duda, no será hasta finales de los años 60 cuando realmente se pueda hablar de una verdadera racionalización y sistemática del uso de los ordenadores aplicados al tratamiento de la información jurídica con los escritos de Mario G. Losano y el surgimiento del concepto de *Iuscibernética*<sup>3</sup>. La Iuscibernética retoma los principios de la cibernética de Wiener en su aplicación al mundo del Derecho, tratando de analizar la lógica del Derecho a fin de introducirla en aquellas aplicaciones informáticas capaces de ser empleadas para el desarrollo de la actividad jurídica, tanto en el campo normativo como el jurisprudencial o doctrinal. De los distintos acercamientos que Losano realiza a este nuevo concepto será el del uso y manejo del ordenador el que derive en el concepto actual de *Informática jurídica*, cuyo objetivo principal, establecido por el propio autor, consiste en “crear programas específicos que permitan memorizar en el ordenador cierto tipo de documentación, así como recuperarla en lo sucesivo seleccionándola por temas”<sup>4</sup>. De este modo, la informática jurídica es considerada por todos los especialistas como una de las partes constituyentes de la iuscibernética y cuya finalidad, eminentemente utilitarista y pragmática, se centra en el tratamiento de la información jurídica por medio las técnicas y procedimientos de la informática para poder de este modo ser recuperada y suministrada de forma rápida y eficaz.

En resumen, nos estaremos refiriendo a la informática jurídica cuando el jurista (juez, secretario, oficial, abogado en el ejercicio de la profesión, académico, etc.) o el documentalista hagan uso de las tecnologías informáticas para procesar, automatizar y sistematizar la información jurídica, en cualesquiera de los niveles en las que se divide ésta (legislación, jurisprudencia y doctrina). Hoy en día, por tanto, la expresión “informática jurídica” se emplea para designar de forma general a todo uso del ordenador dentro del campo del Derecho.

Sin embargo, la simple aplicación de la informática al Derecho no constituye por sí sola una solución a los problemas del control y el acceso a la información jurídica. Esta circunstancia fue prontamente puesta de manifiesto con las primeras experiencias en la aplicación de la informática a los textos jurídicos para la construcción de bases de datos pues se observó claramente que la computadora no era capaz de realizar simples

---

<sup>1</sup> Miguel López-Muñiz recoge en su manual sobre informática jurídica documental las opiniones de Wiener al respecto de este tema y plasmadas en la obra *Cybernetix*. Wiener comenta aquí, entre otras cosas, que “La técnica de interpretación de los precedentes jurídicos debe ser tal que un abogado pueda no solamente saber lo que ha dicho un tribunal, sino también prever con una fuerte probabilidad lo que va a decir ese tribunal. Los problemas del Derecho tienen relación con la comunicación y la cibernética, en el sentido de tratarse de problemas de control regular y dependiendo de ciertas situaciones críticas”. N. Wiener. *Op cit.*, en Miguel López-Muñiz. *Informática Jurídica Documental*. Madrid: Díaz de Santos, 1984, p. 15.

<sup>2</sup> Lee Loevinger. “Jurimetrics. The Next Step Forward”. *Minnesota Law Review*, n° 33, 1949, pp. 455 y ss.

<sup>3</sup> El título exacto de la obra publicada en Turín por este autor en 1969 es *Giuscibernetica. Macchine e modelli cibernetici nel diritto*.

<sup>4</sup> Mario G. Losano. *Introducción a la Informática Jurídica*. Palma de Mallorca: Facultad de Derecho, Universidad de Palma de Mallorca, 1982, p. 37.

operaciones de recuperación de información almacenada. Este hecho fue debido principalmente a la naturaleza del lenguaje natural, impreciso y ambiguo, en el que se expresa el lenguaje jurídico, motivo por el cual se producía el fenómeno del *ruido* así como el del *silencio* en la recuperación de los datos. Como señala R. A. Guibourg, tanto el texto como el título de los documentos jurídicos se componen de expresiones lingüísticas y, por lo tanto, se hallan sujetos a las imprecisiones propias del lenguaje natural; esto es, a la ambigüedad, la vaguedad, las sinonimias, etc.<sup>5</sup>

La respuesta de los juristas fue hacer uso de las técnicas y herramientas que ofrecía la ciencia de la Documentación para el tratamiento de los textos y, en especial, para la construcción de un lenguaje jurídico controlado en donde se da prioridad al estudio y manejo de la semántica y sintaxis de dicho lenguaje para la representación y el procesamiento de los significados presentes en el documento jurídico.

La Documentación tiene una estrecha relación con el Derecho desde su mismo nacimiento<sup>6</sup>. Su surgimiento como disciplina científica tiene sus orígenes a finales del siglo XIX de la mano de dos abogados belgas, Paul Otlet y Henri La Fontaine, con el intento de creación de un inmenso catálogo bibliográfico denominado *Répertoire Bibliographique Universel* (RBU) dentro del seno del *International Institute of Bibliography* (IIB), institución que años más tarde pasaría a tomar el nombre definitivo y actual de *International Federation for Information and Documentation* (FID). El auxilio que esta ciencia proporcionará a los profesionales del Derecho va a derivarse de la misma raíz del asunto, esto es, va a permitir obtener un mayor y mejor conocimiento del objeto mismo del Derecho, el documento jurídico. La Documentación va a proporcionar una serie de técnicas y herramientas al profesional del Derecho que le van a permitir, entre otras muchas cosas, la recuperación lo más exacta posible de la información necesaria para desarrollar su actividad profesional o investigadora.

Así, los documentos jurídicos serán tratados según unas técnicas específicas procedentes de la documentación científica (análisis formal y de contenido, construcción de lenguajes controlados y codificados, etc.), donde un fondo determinado de información jurídica será analizado en sus diferentes niveles para permitir un correcto almacenamiento en la memoria física del ordenador y que pueda ser adecuadamente consultado y recuperado.

La combinación de los saberes y técnicas procedentes de estas tres ciencias dará lugar a una nueva rama de la informática jurídica, la *informática jurídica documental*, dedicada al análisis de los textos jurídicos para la construcción de una serie de herramientas y productos informáticos que permitan obtener de modo más ágil, rápido y preciso que el establecido por los métodos tradicionales, información de carácter jurídico, tanto legislativa como jurisprudencial y/o de doctrina.

## **Las bases de datos jurídicas convencionales**

En cualquier caso, y para los efectos de esta comunicación, el verdadero interés de la informática jurídica documental radica en su capacidad para, a través de las técnicas anteriormente descritas, ser capaz de tratar adecuadamente los textos jurídicos y generar una serie de productos secundarios capaces de ser interrogados satisfactoriamente y

---

<sup>5</sup> Ricardo A. Guibourg, Jorge O. Alende, Elena M. Campanella. *Manual de informática jurídica*. Buenos Aires: Astrea, 1996, p. 80.

<sup>6</sup> Existen innumerables manuales y artículos sobre la ciencia de la Documentación en donde se explica su nacimiento. Véase, por ejemplo, W. Boyd Rayward. "The Origins of Information Science and the International Institute of Bibliography/International Federation for Information and Documentation (FID). *Journal of the American Society for Information Science*, v. 48, n° 2, 1997, pp. 289-300.

producir los resultados esperados por parte de los juristas. Aunque son múltiples los productos que las tareas del análisis documental producen (listados de títulos, de materias, de autores, boletines de resúmenes, etc.), el producto estrella dentro del campo de la informática jurídica lo constituyen las bases de datos, siendo el eje central de la informática jurídica documental. Tal es la importancia de las bases de datos jurídicas y el correcto control de la información contenida en las mismas que, según S. C. de Cartolano, el área documental constituye para el sistema de Derecho de cualquier nación uno de los pilares que sostienen su avance y progreso, debido a la capacidad que le proporciona al jurista para la toma de las decisiones jurídicas oportunas<sup>7</sup>.

El desarrollo de las bases de datos jurídicas ha estado marcado por los avances tecnológicos en este campo concreto de la informática. En cualquier caso, la opción tradicional para la construcción de estos productos informáticos se ha venido basando en la abstracción documental de los textos jurídicos, extrayendo de ellos la información sustancial y estructurando ésta en una serie de campos de interés para su recuperación. Con ello se crean una serie de registros que dan soporte a los modelos relacionales o documentales de la base de datos elaborada. Al amparo de este *modus operandi* han venido surgiendo en estas décadas pasadas un gran número de bases de datos jurídicas en todo el mundo, donde en nuestro país destaca la labor llevada a cabo por editoriales tanto públicas como privadas<sup>8</sup>.

Tal es el caso dentro del primer grupo de las bases de datos jurídicas producidas por el Boletín Oficial del Estado (BOE)<sup>9</sup>, tanto de legislación (Iberlex, Iberlex-UE, MAP-Lexter) como de jurisprudencia (Jurisprudencia Constitucional), todas ellas bases de datos documentales con software de recuperación de información basado en el sistema *Knosys* de la compañía Micronet. En la actualidad, estas bases de datos se distribuyen tanto en soporte óptico (CD-ROM) como en línea.

Entre los editores del sector privado destacan los desarrollos producidos por Aranzadi<sup>10</sup>, sin lugar a dudas la compañía editorial del sector privado en materia de legislación y jurisprudencia más importante de España y la que de forma más decidida ha venido invirtiendo en procesos de tecnificación y adaptación a las nuevas formas de acceso a la información jurídica. Las bases de datos jurídicas producidas por esta institución son ciertamente exhaustivas, recogiendo la práctica totalidad de sentencias y disposiciones a texto completo. Al igual que su versión en línea, recoge información sobre legislación y jurisprudencia en cuatro bases de datos distintas pero que pueden ser cargadas desde un mismo interfaz: Legislación (1930 en adelante), Legislación Autonómica (1978 en adelante para Leyes Autonómicas, Reales Decretos, etc., y desde 1992 para Decretos, Órdenes, etc.), Convenios Colectivos (1979 en adelante), Tribunales Superiores (1979 en adelante) y la de TSJ, AN, TEAC y AP para sentencias de tribunales inferiores (1992 en adelante). Esta empresa fue absorbida en el año 1999 por la compañía editora americana *Thomson*

---

<sup>7</sup> Silvia C. de Cartolano. "Bancos de datos jurídicos". En: *Informática y Derecho*, n° 5, 1994, p. 732.

<sup>8</sup> Para una mayor profundidad en este tema recomendamos la lectura de la obra de Jorge Páez Mañá. *Bases de datos jurídicos: Características, contenido, desarrollo y marco legal*. Madrid: Centro de Información y Documentación Científica, CSIC, 1994.

<sup>9</sup> Toda la información relativa a las diversas bases de datos producidas y distribuidas por el BOE se encuentra disponible en la dirección [http://www.boe.es/bases\\_datos/basesDatos.htm](http://www.boe.es/bases_datos/basesDatos.htm)

<sup>10</sup> La información de los productos comercializados por esta empresa se encuentra disponible en la dirección <http://www.aranzadi.es/>

*Corporation*<sup>11</sup>, convirtiéndose de este modo en el mayor grupo privado de difusión de información jurídica de nuestro país. Sus bases de datos son distribuidas tanto en soporte óptico (CD-ROM y DVD) como a través de un acceso en línea. Hay que destacar también otras compañías españolas, tales como El Derecho Editores<sup>12</sup> o Distribuciones de La Ley<sup>13</sup>.

## La aplicación de los lenguajes de marcado

Sin embargo, frente a estos modelos convencionales para el tratamiento informático y difusión de la información jurídica surge en la década de los 80 una corriente científica crítica, tendente al uso de marcas descriptivas dentro del propio texto de los documentos capaces de caracterizar adecuadamente sus elementos constitutivos, tanto estructurales como puramente semánticos. Con este modo de proceder, no es necesario someter al documento jurídico original a una fase posterior de abstracción documental pues éste incorpora *ab initio* tanto el texto propiamente dicho como las marcas que caracterizan su descripción estructural y semántica, estando, por tanto, ya preparado para su procesamiento por parte de los sistemas de gestión de bases de datos así como para la asignación de un formato o estilo de presentación en uno o varios dispositivos electrónicos de salida (pantalla del ordenador, impresora, dispositivo auditivo, etc.).

## El marcado con SGML

El empleo de los lenguajes de marcado descriptivo tiene sus orígenes a mediados de la década de los 60 en el seno de la *Graphic Communications Association* (GCA) con los trabajos e investigaciones llevados a cabo por William Tunnicliffe y otros colaboradores, orientados al empleo de identificadores genéricos dentro de los textos electrónicos capaces de describir la semántica de los elementos marcados dentro del texto. Pero sin duda, será la aparición a mediados de la década de los 80 del metalenguaje SGML (*Standard Generalized Markup Language*) el hito que marcaría el verdadero inicio y expansión en el empleo de los lenguajes de marcado entre los investigadores y profesionales de los diversos campos del saber humano. Basado en las investigaciones llevadas a cabo por Charles F. Goldfarb, Edward J. Mosher y Ray Lorie dentro de los laboratorios de la compañía IBM, sus inicios están estrechamente relacionados con el ámbito jurídico<sup>14</sup> pues el lenguaje primitivo que se desarrolló aquí, el TDL (*Text Description Language*), fue utilizado inicialmente para el desarrollo de un modelo de sistema de procesamiento integral de textos electrónicos en el ámbito jurídico (edición de texto, almacenamiento y recuperación documental y composición del texto).

SGML es un estándar de alcance internacional (Norma ISO 8879:1986) con una doble funcionalidad: un lenguaje de marcado válido tanto para el procesamiento de textos electrónicos como de utilidad para la gestión de la información en entornos corporativos automatizados. SGML no es en sí un lenguaje específico de marcado de documentos electrónicos sino, más bien, un *metalenguaje*, esto es, un lenguaje que permite crear múltiples

---

<sup>11</sup> <http://www.thomson.com/>

<sup>12</sup> <http://www.elderecho.com/>

<sup>13</sup> <http://www.laley.net/>

<sup>14</sup> De hecho, el principal investigador de este equipo de trabajo, Charles F. Goldfarb, es graduado en Derecho por la Universidad de Harvard (1964).

lenguajes de marcado descriptivo. Por tanto, SGML define dos tipos de sintaxis para un lenguaje de marcado, proporcionando un mecanismo normalizado para la generación de una familia de lenguajes de marcado descriptivo que pueden ser utilizados, principalmente, para la descripción de la estructura de múltiples tipos de documentos electrónicos. Estas sintaxis son: una abstracta, utilizada para declarar las reglas que definen la inserción de marcas descriptivas en los documentos sin tener en cuenta los caracteres específicos utilizados para representar dicho marcado; y otra concreta, que detalla aspectos tales como los caracteres que van a ser utilizados para los delimitadores de marcado, las cantidades, los nombres utilizados en las declaraciones de marcado, etc<sup>15</sup>. Como señala S. Adler, el marcado en SGML está diseñado principalmente para identificar objetos dentro del documento de acuerdo a su función o propósito, y no, dado que es un lenguaje descriptivo, para detallar de las características tipográficas de dichos objetos<sup>16</sup>.

La estructuración de los documentos electrónicos resulta un factor clave en el modelo SGML pues con ello se obtienen las ventajas derivadas de los modelos basados en la estructuración de la información (accesibilidad a partes concretas del documento, validación de los contenidos sujetos a una determinada estructura y la reutilización de contenidos del documento para múltiples propósitos, entre otras), añadiéndose todas aquellas que se derivan de los formatos estandarizados para el intercambio de textos a través de redes informáticas (formato independiente para el intercambio de datos electrónicos entre las organizaciones, longevidad del formato y respaldo por parte de la industria, entre otras)<sup>17</sup>.

Si hemos señalado anteriormente el hecho de ser SGML un metalenguaje de propósito general ello significa que éste debe incorporar un mecanismo capaz de definir lenguajes específicos de marcado; esto se consigue a través de la denominada DTD (*Document Type Definition*). La DTD definirá la estructura lógica genérica que han de tener los documentos electrónicos que se ajusten a ella, detallando para ello una serie de propiedades que han de contemplar los elementos que se integran dentro del marcado del documento electrónico, tales como los nombres de los elementos que están permitidos, la frecuencia de aparición de cada uno de los elementos, el orden de aparición de éstos, el contenido permitido para cada uno de ellos, los atributos que pueden definir algunas características particulares de algunos elementos, y otras cuestiones similares<sup>18</sup>.

La aplicación del lenguaje SGML para la construcción de documentos electrónicos en el ámbito jurídico fue paradójicamente tardía, pues no es hasta pasados algunos años de la década de los 90 cuando se puede hablar de la aparición de proyectos de envergadura para el tratamiento de documentos de carácter legal.

Como señala M. Blaquier, las primeras aplicaciones del estándar SGML en el campo jurídico fueron ligadas a la publicación de documentos normativos por diversas editoriales especializadas, tanto del sector privado como del público<sup>19</sup>, que emplearon inicialmente el

---

<sup>15</sup> Charles F. Goldfarb, Yuri Rubinsky. *The SGML Handbook*. [1ª ed., 4ª reimp.] New York [etc.]: Oxford University Press, 1990, pp.135-136.

<sup>16</sup> Sharon C. Adler. "The Birth of a Standard". *Journal of the American Society for Information Science*, v. 43, nº 8, 1992, p. 556.

<sup>17</sup> *Which is for Me? Structured Documents or SGML?* [documento HTML]. SGML Associates, 1996. Disponible en <http://www.mcs.net/~dken/struct.htm> (consultado el 5 de noviembre de 2001).

<sup>18</sup> Eric van Herwijnen. *Practical SGML* [2ª ed.]. Boston [etc.]: Kluwer Academic Publishers, 1994, pp. 32-33.

<sup>19</sup> Marta Blaquier Ascaño. "Aplicaciones del SGML a la Información Jurídica". *Informática y Derecho*, nº 12-15, 1996, p. 1489.

SGML como medio idóneo para una posterior publicación de la información legal a través de los tradicionales medios impresos. Entre las empresas editoras privadas cabe destacar por su pronta utilización para la confección, almacenamiento y publicación de normas legales a la casa editora *Kluwer*<sup>20</sup> en Holanda y la *Thompson Legal Publishing*<sup>21</sup>, que creó el *Research Institute of America and Carswell* como subdivisión para la investigación y uso de SGML en la publicación de documentos jurídicos. Pero es sin duda en el sector de los editores públicos donde se han llevado a cabo las experiencias de trabajo con SGML más importantes. Así, cabe destacar las iniciativas emprendidas a mediados de los 90 por la *Government Printing Office*<sup>22</sup> de los Estados Unidos para la transferencia de los documentos producidos por el Congreso americano a un sistema basado en el lenguaje SGML para, de este modo, facilitar las tareas de búsqueda y localización de información de carácter oficial<sup>23</sup>; la *Her Majesty's Stationery Office*<sup>24</sup>, imprenta y editora oficial del gobierno británico, donde desde comienzos de 1987 toda la legislación emanada del parlamento británico era preparada y codificada en SGML de acuerdo a una compleja DTD, con la doble intencionalidad de que el marcado de los textos legales fuese de utilidad tanto para los sistemas de impresión y publicación como para el sistema de almacenamiento y gestión de la información en las correspondientes bases de datos, de tal forma que una nueva ley pudiese aparecer al mismo tiempo publicada en el boletín impreso tradicional así como en formato electrónico<sup>25</sup>; o, finalmente, la *Official Publication Office*<sup>26</sup> de las Comunidades Europeas, que desarrolló a mediados de la década de los 80 un sistema denominado FORMEX (*Formalised Exchange of Electronic Documents*) con el fin de facilitar la producción de las publicaciones oficiales de la Comisión, en especial el Diario Oficial, así como la mejora en los sistemas de almacenamiento y recuperación de estos documentos en formato electrónico. Este sistema ha venido evolucionando notablemente desde su lanzamiento, con diversas versiones mejoradas, constituyendo hoy en día uno de los ejemplos más atractivos del empleo del SGML para el tratamiento y publicación de documentos de carácter legal<sup>27</sup>. Dentro del ámbito jurídico resulta destacable por su importancia y uso extendido entre los profesionales del derecho comunitario el sistema automatizado de documentación para la información legal emanada de la UE, conocido como CELEX

---

<sup>20</sup> <http://www.kluwer.nl/>

<sup>21</sup> <http://www.thompson.com/>

<sup>22</sup> <http://www.access.gpo.gov/>

<sup>23</sup> *SGML in the Law* [documento HTML]. Los Angeles, CA: University of South California, Law School, 12 de septiembre de 1995. Disponible en <http://www.usc.edu/dept/law-lib/rlg/sgml.html> (consultado el 11 de abril de 2002).

<sup>24</sup> <http://www.hmso.gov.uk/>

<sup>25</sup> R. Stutely. "HMSO's Database of Legal Text". *SGML Users' Group Bulletin*, v. 2, n° 2, 1987. Citado en David Barron. *Why use SGML?* Electronic Publishing, v. 2, n° 1, April 1989, pp. 3-24. Disponible en <http://cajun.cs.nott.ac.uk/compsci/epo/papers/volume2/issue1/epdxb021.pdf> (consultado el 11 de abril de 2002).

<sup>26</sup> <http://www.eur-op.eu.int/>

<sup>27</sup> <http://www.eur-op.eu.int/general/fr/b4.htm>. Existe igualmente una publicación oficial impresa sobre este desarrollo elaborada en 1985 por la propia OPOCE: Office for Official Publications of the European Communities. *FORMEX – Formalized Exchange of Electronic Publications*. Luxembourg: OPOCE, 1985.

(*Comunitatis Europae LEX*)<sup>28</sup>. Esta magna base de datos, producida como un sistema interinstitucional, es alimentada a través del sistema de entrada de datos de diferentes fuentes, denominado ALICE (*ALImentation CElex*), que está basado por completo en el estándar SGML, utilizado no sólo para la definición del formato de entrada de los documentos sino también para el formato básico bajo el cual los documentos son almacenados para el intercambio entre el sistema de producción y de difusión de las distintas bases de datos<sup>29</sup>.

En el ámbito de la administración de justicia también encontramos notables ejemplos de aplicación del estándar SGML al tratamiento de la documentación producida. Algunos de los proyectos más destacables en este campo los podemos encontrar en Estados Unidos, Canadá y Europa.

Para el caso de los Estados Unidos, encontramos ejemplos como el del *Delaware Superior Court* con el lanzamiento del proyecto CLAD (*Complex Litigation Automated Docketing*)<sup>30</sup> con la finalidad de facilitar la gestión de litigios de mayor complejidad que se producían en dicho tribunal estatal, demostrándose la viabilidad de esta tecnología y la gran reducción de costes económicos que llevaba aparejada. Igualmente interesante resulta el proyecto del *Snake River Basin Adjudication* (SRBA) – *District Court* del estado de Idaho<sup>31</sup>, que surge también para dar respuesta a una situación jurídica de gran complejidad: la gran cantidad de reclamaciones tramitadas en este tribunal debido a la repartición de las aguas de la cuenca del río Snake. El proyecto más ambicioso, el *Utah Electronic Filing Project*, emprendido por la *Utah Administrative Office of the Courts*<sup>32</sup>, ha sido citado con frecuencia como caso ejemplar de uso del SGML en la administración de justicia norteamericana. Se basa en la utilización de tecnologías informáticas para agilizar todos los procesos administrativos y judiciales que se dan en estas instituciones; el estándar SGML ha jugado un papel crucial en su desarrollo. Finalmente conviene destacar la labor llevada a cabo por el Tribunal Supremo<sup>33</sup> de este país con el desarrollo y puesta en marcha del proyecto HERMES para el tratamiento y distribución de las resoluciones, que arrancó en mayo de 1990 con la participación conjunta del citado Tribunal, la *Government Printing Office* y algunas universidades.

El caso canadiense no es menos atractivo pues se han dado notables ejemplos del uso del SGML para el tratamiento de la documentación jurídica. El más importante es el del Tribunal Supremo de Canadá<sup>34</sup>, que decidió poner en marcha un proyecto, en colaboración con el *Centre de Recherche en Droit Public* (CRDP)<sup>35</sup> de la Universidad de Montreal, para poner

---

<sup>28</sup> Más información, y acceso a la consulta de esta base de datos, en <http://europa.eu.int/celex/>

<sup>29</sup> J. Marín Navarro, P. E. Alevantis. "Alice in the Wonderland of SGML: streaming text entry in the CELEX databases". *The Electronic Library*, v. 9, n° 3, June 1991, p. 155.

<sup>30</sup> Una profunda descripción histórica y funcional de este proyecto se encuentra en la comunicación de Robert J. Katzenstein. "CLAD – Delaware's Paperless Docket" [documento HTML]. En: *National Court Technology Conference (CTC4)* (4°. 1994. Washington D.C.). Washington D.C.: National Center for State Courts, October 1994. Disponible en [http://www.ncsc.dni.us/ncsc/ctc4/articles/elec\\_b.htm](http://www.ncsc.dni.us/ncsc/ctc4/articles/elec_b.htm) (consultado el 11 de abril de 2002).

<sup>31</sup> La información oficial sobre este proyecto está en <http://www.idwr.state.id.us/info/water/srba/main.htm>

<sup>32</sup> <http://courtlink.utcourts.gov/>

<sup>33</sup> <http://www.supremecourtus.gov/>

<sup>34</sup> <http://www.scc-csc.gc.ca/>

<sup>35</sup> <http://www.crdp.umontreal.ca/>



disposición del público sus principales colecciones de material jurídico, empleando tecnologías informáticas de gran alcance y distribución, como es el caso de Internet. Los trabajos iniciales comenzaron en la primavera de 1994 y, ante la necesidad de establecer un sistema de tratamiento documental que fuese riguroso y preciso, se consideró a SGML el formato más adecuado.

Dentro del continente europeo destacan dos proyectos de investigación aplicada a la documentación jurídica llevados a cabo en Finlandia y Suecia. El primero de ellos, el denominado *Raske Project*<sup>36</sup>, encabezado por el *Department of Computer Science and Information Systems* de la Universidad de Jyväskylä, junto con el Parlamento de Finlandia y diversos ministerios de este país, se puso en marcha en 1994. Sobre la base de un profundo análisis en la tipología documental producida por estas instituciones, se desarrollaron tres DTD SGML para tres tipos de documentos (proyectos de ley del gobierno, diligencias realizadas por el Parlamento al gobierno, y las réplicas de éste) que fueron además definidas para generar documentos en diversos idiomas. El primer prototipo, utilizando unas definiciones preliminares de tipo de documento SGML, fue evaluado en 1995<sup>37</sup>. La complejidad de este proyecto ha sido extrema pues siempre ha tenido un planteamiento global de gestión de la información producida por estas instituciones y de mejora de los canales de producción y de comunicación de la misma. Finalmente, destacar por su importancia el proyecto sueco *Corpus Legis Project*<sup>38</sup>, puesto en marcha por el *Swedish Law & Informatics Research Institute* de la *Faculty of Law* con la colaboración del *Department of Computational Linguistics*, ambos de la Universidad de Estocolmo. Iniciado en 1994 y finalizado en 1998, tenía como objetivo principal el establecimiento de mecanismos que mejorasen el acceso y la recuperación de información jurídica, para dar respuesta a su crecimiento exponencial y a la necesidad de armonización europea en esta materia. Al igual que en los anteriores proyectos, en éste la piedra angular era la utilización de sistemas informáticos basados en el estándar SGML, por lo que una gran parte del proyecto se basó en el análisis exhaustivo de los corpus textuales de los documentos generados por las distintas instituciones afectadas.

## El mercado con HTML

Pero será sin duda la popularización de la red Internet y de la *World Wide Web*, el principal motor que impulse la utilización de los lenguajes de marcado de texto aplicados a la documentación jurídica. Internet y sus servicios característicos han venido a potenciar de manera espectacular el concepto de la llamada *Sociedad de la Información*, la cual implica, entre otras muchas cosas, una sociedad en donde el uso de la información electrónica y las comunicaciones se convierten en la pieza central que mueve el desarrollo de las sociedades modernas. Dentro de toda esa maraña de información electrónica que circula por Internet, el acceso a la información jurídica por parte de los ciudadanos juega un papel determinante en el desarrollo de este modelo de Sociedad.

Sin entrar en el debate sobre la conveniencia de que las distintas administraciones públicas empleen las nuevas tecnologías de Internet para poner a disposición del ciudadano en general la información que de ellas emana, que en el caso jurídico han venido exponiendo

---

<sup>36</sup> <http://www.cs.jyu.fi/~raske/>

<sup>37</sup> Jarmo Kiuru, Airi Salminen, Su-Shing Chen. *Digital Archiving of Legislative Documents: A Case Study of Finland* [documento HTML]. Washington D.C.: NSF Workshop on Data Archival and Information Preservation, 1999. Disponible en <http://cecssrv1.cecs.missouri.edu/DA+IPpapers/finland.html> (consultado el 11 de abril de 2002).

<sup>38</sup> <http://www.juridicum.su.se/iri/corpus/>

diversos profesionales del Derecho, como Thomas R. Bruce, del Instituto Jurídico de la Escuela de Derecho de la Universidad norteamericana de Cornell<sup>39</sup>, lo que sí resulta un hecho cierto en nuestros días es la especial sensibilización que gobiernos centrales o autonómicos en muchos países están demostrando en esta materia, desarrollando en fechas recientes auténticos *portales web* con información sobre las distintas administraciones y servicios públicos, ofreciendo al ciudadano un espacio centralizado para la localización de la información deseada o, en muchos casos, para establecer un nuevo mecanismo de comunicación y relación (electrónica) con dichas Administraciones<sup>40</sup>.

Centrándonos en el caso concreto del suministro de información jurídica en el espacio electrónico de la WWW, S. Erdelez y S. O'Hare señalan que el uso de la Web constituye desde su propia filosofía de construcción una manifestación perfecta de lo que es en realidad el entramado de información y datos relacionados que caracteriza a los documentos jurídicos<sup>41</sup>. En los medios impresos tradicionales, e incluso en los primeros sistemas de acceso a textos electrónicos, el jurista tenía que ir de un modo lineal de un texto a otro para obtener una información completa de lo que deseaba localizar. Sin embargo, con la concepción en el modo de funcionamiento de la Web y la materialización en la construcción de documentos hipertextuales basado en el lenguaje HTML (*HyperText Markup Language*), los documentos no están ya limitados por las dos dimensiones de la hoja de papel. Este nuevo entorno de consulta y lectura no secuencial de la información permite al usuario moverse o *saltar* desde un concepto existente en un documento a otro concepto relacionado, dentro del mismo documento o en otro distinto, formando una maraña de enlaces que ponen en relación informaciones y documentos vinculados lógicamente. En el caso de los documentos jurídicos parece absolutamente clara esta funcionalidad pues son numerosas las referencias que se dan dentro de una norma legal o sentencia judicial a otras normas y sentencias anteriores fundamentando jurídicamente su desarrollo.

Esta característica fundamental de los documentos jurídicos es puesta de manifiesto por J. T. Nogales y M. C. Arellano, al señalar que las relaciones que explícitamente se establecen en el texto de los documentos jurídicos a otros hacen que éstos no puedan ser entendidos como entes aislados sino, más bien, como una unidad de contenido, un sistema global (el ordenamiento jurídico), compuesto por diferentes elementos con significado propio<sup>42</sup>.

Con respecto a las ventajas que la tecnología Web y los lenguajes de marcado de los documentos que residen en este espacio de publicación aportan a los documentos jurídicos

---

<sup>39</sup> Véanse, por ejemplo, los artículo "Public Legal Information: Focus and Future" [documento HTML]. *The Journal of Information, Law and Technology* (JILT), n° 1, 2000. Disponible en <http://elj.warwick.ac.uk/jilt/00-1/bruce.html> (consultado el 11 de abril de 2002) o, de igual interés, *Some Thoughts on the Constitution of Public Legal Information Providers* [documento HTML]. Ithaca, NY: Cornell Law School, Cornell University, 2000. Disponible en <http://www4.law.cornell.edu/working-papers/open/bruce/warwick.html> (consultado el 11 de abril de 2002).

<sup>40</sup> Sirva como ejemplo lo sucedido en España con el *portal* que la Administración General del Estado ha puesto en marcha en fechas recientes. Aunque aún no está desarrollado de forma completa, pues los servicios que ofrece son menores que los de sistemas similares de otros países de nuestro entorno, desde aquí es posible acceder a gran parte de la información que se suministra en diversos organismos y dependencias de la Administración Central, así como a la consulta de un gran número de bases de datos de todo tipo. Para más información véase <http://www.administracion.es/>

<sup>41</sup> Sanda Erdelez, Sheila O'Hare. "Legal Informatics: Applications of Information Technology in Law". En: Martha E. Williams (ed.). *Annual Review of Information Science and Technology*, v. 32. Medford, NJ: Information Today, 1997, p. 389.

<sup>42</sup> J. Tomás Nogales Flores, M. Carmen Arellano Pardo. "La organización hipertextual de textos legislativos con HTML y XML: una necesidad y las soluciones de presente y futuro". En: *Jornadas Españolas de Documentación* (7ª. 2000. Bilbao). Bilbao: Universidad del País Vasco, 2000, p. 180.

y, por extensión, a la construcción y diseminación de esta información contenida en bases de datos legislativas y jurisprudenciales, estos mismos autores señalan las siguientes<sup>43</sup>:

- *Tecnología de gran implantación.* Los navegadores o *browsers* de web están implantados en cualquier ordenador con acceso a Internet y HTML se ha convertido en un formato de intercambio universal de documentos textuales. Igualmente, existen numerosas aplicaciones, muchas de ellas gratuitas, tanto para el establecimiento de servicios de información en este espacio electrónico como para el acceso a los mismos.
- *Documentos de cualquier tamaño y con integración de medios diversos.* Además de texto en lenguaje natural, los documentos pueden integrar otros medios como tablas, imágenes, gráficos o diagramas (presentes ocasionalmente en textos normativos), sonido, vídeo, etc.
- *Capacidad hipertextual.* Las referencias explícitas que aparecen en el texto (tan frecuentes e importantes en los documentos jurídicos), tanto las internas al documento como las externas a otros documentos, pueden ser integradas en el propio texto. Incluso las implícitas pueden ser determinadas por especialistas e incluidas en los documentos.
- *Diversidad de soportes y medios de difusión para una misma base de datos.* Sin establecer modificación alguna, el conjunto de documentos hipertextuales que conforma la base de datos puede ser difundido a través de Internet por medio de un servidor web o bien, en la forma tangible del CD-ROM, DVD u otros soportes ópticos y magnéticos; sobre él además es posible integrar un motor de búsqueda que permita realizar las funciones clásicas de búsqueda de información, en especial si se utilizan las últimas tecnologías aplicadas a la Web, caso del metalenguaje XML.

## El marcado con XML

XML (*eXtensible Markup Language*)<sup>44</sup> es el nuevo metalenguaje desarrollado por el *World Wide Web Consortium (W3C)*<sup>45</sup>, reducción y adaptación de SGML al espacio de publicación de Internet. Establecido como recomendación oficial de esta organización en febrero de 1998<sup>46</sup>, permite a los usuarios, al igual que SGML, crear sus propios lenguajes o vocabularios de marcado de textos electrónicos a través del mecanismo de la DTD, definiendo las estructuras lógicas subyacentes a cada tipo documental tratado, así como una asignación semántica de cada una de las piezas u objetos existentes en dichas estructuras.

Al amparo de este metalenguaje de marcado han venido surgiendo en estos últimos años otros estándares que lo complementan con diversas capacidades, como XLink (*Extensible Linking Language*) para la creación de enlaces hipertextuales avanzados dentro de los documentos XML, XML Namespaces para la integración de diferentes vocabularios, XSL

---

<sup>43</sup> J. T. Nogales Flores, M. C. Arellano Pardo. *Op. cit.*, p. 181.

<sup>44</sup> Toda la información oficial relativa al Lenguaje Extensible de Marcado (XML) se encuentra ubicada en el sitio Web del *W3 Consortium*, en la dirección <http://www.w3.org/XML/>

<sup>45</sup> <http://www.w3.org/>

<sup>46</sup> World Wide Web Consortium. *Extensible Markup Language (XML) 1.0. W3C Recommendation* [documento HTML]. Tim Bray, Jean Paoli, C.M. Sperberg-McQueen (eds.). W3C, February 10, 1998. Disponible en <http://www.w3.org/TR/1998/REC-xml-19980210> (consultado el 11 de abril de 2002). Existe una revisión o segunda edición del mismo de octubre de 2000: W3C. *Extensible Markup Language (XML) 1.0 (Second Edition). W3C Recommendation* [documento HTML]. Tim Bray, Jean Paoli, C.M. Sperberg-McQueen, Eve Maler (eds.). W3C, October 6, 2000. Disponible en <http://www.w3.org/TR/2000/REC-xml-20001006> (consultado el 11 de abril de 2002).

(*Extensible Stylesheet Language*) para la asignación de un formato o estilo de presentación a estos documentos, XML Schema para la definición formal de estructuras documentales y tipos de datos complejos o RDF (*Resource Description Framework*) para la asignación de metainformación a los objetos electrónicos que pueblan la red.

Son innumerables las ventajas que el metalenguaje XML aporta a la nueva Web que se desea construir, a esa segunda generación de este espacio electrónico definido por J. Bosak y T. Bray a principios de 1999<sup>47</sup>, entre las que pueden destacarse las siguientes: es una tecnología abierta y, por tanto, no propietaria, independiente de la plataforma y del sistema operativo; es un estándar de alcance internacional basado en el sistema Unicode de codificación de caracteres; es una tecnología sencilla de utilizar y fácil de implantar; tiene una gran potencia para la construcción de vocabularios de marcado aplicables a cualquier tipo de documento; se orienta al tratamiento, transmisión o intercambio de todo tipo de documentos o datos; permite reutilizar textos y datos ya existentes en otros documentos para la confección de nuevos documentos; y, especialmente, aporta potentes mecanismos para la búsqueda y recuperación de la información.

Desde el primer momento de la creación del XML la comunidad científica y profesional del campo jurídico ha vislumbrado el potencial de esta tecnología, que puede ser una pieza fundamental para el correcto tratamiento de la información jurídica emanada de las distintas instituciones del estado. Son numerosos los ejemplos que podemos encontrar en diferentes países del mundo sobre la aplicación de esta tecnología al tratamiento y difusión de información jurídica, especialmente dentro del ámbito judicial.

Es, sin duda, en los Estados Unidos donde de forma más clara y decidida se han emprendido la mayor parte de los proyectos de investigación con casos muy significativos tanto a escala estatal como federal. Dentro del primer grupo, podemos destacar los tempranos esfuerzos llevados a cabo por el *College of Law at Georgia State University*<sup>48</sup> en colaboración con la Escuela de Negocios de J. Marck Robinson<sup>49</sup> de dicha Universidad, con la puesta en marcha del proyecto *Electronic Court Filing Project (E-CT-Filing)*<sup>50</sup>, dirigido por Winchel "Todd" Vincent y cuya finalidad inicial era analizar los diferentes sistemas existentes para la mejora de los procesos de gestión e intercambio de documentos electrónicos entre los tribunales de justicia del estado citado, siendo el metalenguaje XML el formato de codificación de documentos más destacado. Este grupo de trabajo se une con el grupo de investigadores de la Universidad de Utah, que venían trabajando en el proyecto *Utah Electronic Law & Commerce Partnership (UELCP)*<sup>51</sup>, y con el investigador Gabe Wachob del portal jurídico FindLaw<sup>52</sup> para la colaboración científica entre todas estas instituciones y

---

<sup>47</sup> Jon Bosak, Tim Bray. "XML and the Second-Generation Web" [documento HTML]. *Scientific American*, May 1999. Disponible en <http://www.sciam.com/1999/0599issue/0599bosak.html> (consultado el 11 de abril de 2002).

<sup>48</sup> <http://law.gsu.edu/>

<sup>49</sup> <http://www.cba.gsu.edu/robinson/home.asp>

<sup>50</sup> Toda la información oficial relativa a dicho proyecto se encuentra disponible en la dirección <http://gsulaw.gsu.edu/gsupecp/>

<sup>51</sup> <http://www.uelcp.org/default.cgi>

<sup>52</sup> Se trata de un clásico portal web dedicado íntegramente a la localización de recursos informativos de interés jurídico en Internet. Se estructura en un árbol jerárquico de forma similar a los directorios de recursos (al estilo de Yahoo!), con un primer nivel según el tipo de usuario: profesionales de la actividad jurídica, estudiantes, sector comerciales o de negocios, recursos humanos y público en general. Dispone igualmente de un motor de búsqueda para la localización directa de los recursos. Accesible en <http://www.findlaw.com/>

la revitalización del primitivo grupo de trabajo sobre XML creado un año antes en la Universidad de Utah<sup>53</sup>. El nuevo grupo de trabajo de la Universidad del estado de Georgia y las otras dos instituciones ponen en marcha en noviembre de 1998 un servicio de listas de distribución de correo electrónico para el debate en línea entre los investigadores del país interesados en la materia (en aquellos momentos unos 18 miembros). Este primitivo grupo de discusión y trabajo ha venido siendo considerado como el verdadero inicio de la organización internacional sin ánimo de lucro Legal XML<sup>54</sup>, en la que se integran personas procedentes de todo tipo de instituciones del campo jurídico con el objetivo de desarrollar estándares abiertos y no propietarios de aplicación jurídica basados en el metalenguaje XML. En el seno del grupo de trabajo conocido por el nombre de Court Filing Workgroup<sup>55</sup> se han venido desarrollando diversos vocabularios de marcado relacionados con los documentos y los formatos de intercambio de información electrónica para las aplicaciones informáticas propias de los tribunales de justicia, tales como el *LegalXML Court Filing 1.0 Standard*<sup>56</sup>, de enero de 2000, y el *XML Court Document Draft Standard*<sup>57</sup>, de junio de 2001 aún en fase de borrador.

Resulta igualmente encomiable la labor llevada a cabo durante esos años por Richard Himes dentro del *United States District Court, District of New Mexico*<sup>58</sup>, donde se ha venido trabajando desde mediados de 1998 en el desarrollo conocido como *Extensible Markup Language Court Interface (XCI)*<sup>59</sup> con la finalidad de alentar el desarrollo e investigación de estándares independientes para los sistemas de edición y transmisión electrónica de documentos judiciales.

---

<sup>53</sup> En <http://www.uelcp.org/xml.cgi> hay más información sobre este primitivo pero importante grupo de trabajo, liderado por Brent Israelsen de la Universidad de Utah, que se formó para el debate intelectual y el desarrollo de una o varias DTDs XML normalizadas para la edición e intercambio de documentos electrónicos jurídicos, pero que, desafortunadamente, se quedó en la simple discusión.

<sup>54</sup> <http://www.legalxml.org/>

<sup>55</sup> La información oficial de este grupo de trabajo se encuentra disponible en la dirección <http://www.legalxml.org/CourtFiling/>

<sup>56</sup> La autoría de esta DTD corresponde a Marty Halvorson y Richard Himes, y se encuentra en la dirección [http://www.legalxml.org/DocumentRepository/ProposedStandards/Clear/PS\\_10001/PS\\_10001\\_2000\\_07\\_24.htm](http://www.legalxml.org/DocumentRepository/ProposedStandards/Clear/PS_10001/PS_10001_2000_07_24.htm)

<sup>57</sup> Mohyeddin Abdulazziz, Rolly Chambers, John Messing. *Legal XML Proposed Standard: XML Standards Development Project – XML Court Document 1.0 Draft Standard* [documento HTML]. Legal XML, June 21, 2001. Disponible en [http://www.legalxml.org/appeals/courtdocument/Legal%20XML%20Court%20Document%201.0%20Draft%20DTD\(rev\).htm](http://www.legalxml.org/appeals/courtdocument/Legal%20XML%20Court%20Document%201.0%20Draft%20DTD(rev).htm) (consultado el 23 de agosto de 2001). Éste y otros documentos relacionados (documentos informativos, ficheros de la DTD con cada fecha de revisión, diagramas de las mismas y ejemplos prácticos) se encuentran en <http://www.legalxml.org/appeals/court filingdocument.htm>

<sup>58</sup> <http://www.nmcourt.fed.us/dcdocs/>

<sup>59</sup> Aunque la información relativa a este desarrollo se debería encontrar disponible en su sitio oficial (<http://www.nmcourt.fed.us/xci/xcihome.htm>), algunos documentos electrónicos de importancia han sido eliminados de este servidor. Por ello, recomendamos que para consultar uno de los mejores documentos explicativos del funcionamiento del XCI sea acceda a través de la copia local existente en el servidor Web de la organización OASIS, en <http://www.oasis-open.org/cover/xcispec19990402.html>

El estado de California desarrolló el denominado *Open XML Court Interface (OXCI)*<sup>60</sup>, de similares características al anterior, puesta en marcha en marzo de 2000 con la intención de desarrollar un estándar abierto para un sistema de intercambio de documentos electrónicos de interés judicial en los tribunales de este estado.

Existen, sin embargo, otras organizaciones en los Estados Unidos de carácter supraestatal que están tratando de llevar a cabo la coordinación de todas las iniciativas desarrolladas en los diversos estados con la finalidad de obtener un estándar que sea a la vez común para todos los tribunales de justicia de la nación pero con la suficiente flexibilidad para adecuarse a la variaciones locales que se pudieran producir. Entre este tipo de organizaciones son destacables el *National Center for State Courts (NCSC)*, la *Conference of State Court Administrator (COSCA)*, la *National Association for Court Management (NACM)* y la *American Bar Association (ABA)*<sup>61</sup>.

En el continente europeo destacan una serie de proyectos enmarcados dentro de las políticas generales de información que diversos países han ido poniendo en marcha en estos últimos años. Tal es el caso del Reino Unido con el lanzamiento de un plan de modernización tecnológica de las administraciones públicas a través del denominado programa e-GIF (*e-Government Interoperability Framework*)<sup>62</sup>. Este marco de trabajo establece las vías para la modernización del gobierno y de sus instituciones a través de la integración de los sistemas de información existentes en el sector público y define los requisitos previos considerados esenciales para lograr esta integración tecnológica de los sistemas de información gubernamentales y su disponibilidad por parte del ciudadano a través de la web tomando como base tecnológica el metalenguaje XML. De forma más detallada, este ambicioso marco de trabajo señala tres puntos clave de desarrollo y aplicación de la tecnología XML al e-GIF: en primer lugar, el metalenguaje XML y los esquemas XML para la integración de los datos; en segundo lugar, la aplicación de los estándares UML, RDF y XML para la modelización de los datos y el lenguaje descriptivo; y tercero, el empleo de los estándares XSL, DOM y XML para la presentación de los datos.

Francia es otro de los países europeos que de forma más decidida está apostando por el desarrollo tecnológico e informático de sus administraciones públicas y donde, como expondremos, el desarrollo y aplicación de los estándares XML juega un papel importante. Además de la existencia de diversos foros electrónicos de debate sobre XML, resulta especialmente destacable el *Club XML*<sup>63</sup>, surgido al amparo de las decisiones adoptadas por el *Comité Interministériel à la Réforme de l'Etat* en octubre de 2000, relativas a la utilización de las tecnologías XML para el desarrollo de un amplio repertorio de esquemas XML de aplicación en las diversas administraciones públicas, y que se conforma como un elemento

---

<sup>60</sup> Richard Himes. *Open XML Court Interface (OXCI). Working Draft* [documento PDF]. LegalXML, October 15, 2000. Disponible en <http://www.legalxml.org/California/files/OXCIArchitecture2.pdf> (consultado el 16 de febrero de 2001).

<sup>61</sup> Los URL de estas instituciones son, respectivamente, <http://www.ncsconline.org/>, <http://cosca.ncsc.dni.us/>, <http://www.nacmnet.org/> y <http://www.abanet.org/>

<sup>62</sup> Toda la información oficial relativa a este plan, que desde mayo de 2001 se encuentra en su versión 2, se encuentra disponible en el sitio web del organismo oficial denominado *Office of the e-Envoy* en la dirección <http://www.citu.gov.uk/publications/frameworks/egif2/execsum.htm> La versión 2 del documento oficial que da soporte a este marco estratégico se encuentra disponible en *E-Government Interoperability Framework* [documento PDF]. London: Office of the e-Envoy, May 2001. Disponible en [http://www.e-envoy.gov.uk/publications/frameworks/egif2/e-GIF\\_ver2.pdf](http://www.e-envoy.gov.uk/publications/frameworks/egif2/e-GIF_ver2.pdf) (consultado el 11 de abril de 2002). Este documento electrónico se encuentra disponible en otros formatos (Word, RTF y HTML).

<sup>63</sup> Toda la información oficial relativa a este grupo de investigación y trabajo se encuentra disponible en <http://www.atika.pm.gouv.fr/XML/clubxml/introduction.shtml>

fundamental para el debate e intercambio de ideas entre aquellos profesionales e investigadores del país interesados en el desarrollo de dicho repertorio<sup>64</sup>. Así, y fruto de esta actividad, se han venido poniendo en marcha en estos últimos años varios proyectos de investigación en el desarrollo de esquemas XML y DTD aplicados a la información electrónica de diversos organismos públicos del Estado<sup>65</sup>. Sirva como ejemplo el caso del Consejo Constitucional<sup>66</sup> francés donde se está desarrollando un proyecto de investigación aplicada bajo tecnologías XML, aún en fase de experimentación, que se enmarca dentro de otro proyecto global del estado francés denominado *Service Public d'Accès au Droit* (SPAD)<sup>67</sup> cuya finalidad principal es crear un servicio público de acceso al Derecho que ponga a disposición de todos los ciudadanos, a través de Internet y de forma gratuita, toda la legislación y jurisprudencia emanada de los diversos órganos competentes en la materia. En ello está trabajando esta institución con el desarrollo de una DTD para marcar los documentos jurídicos emanados de esta institución. Existe definida ya una DTD XML<sup>68</sup> para hacer posible esta transmisión de documentos jurídicos del Consejo Constitucional para su difusión a través del futuro servicio SPAD.

En la Unión Europea podemos encontrar un caso concreto de la aplicación de las tecnologías XML al campo de la justicia en la aplicación del Proyecto GTi (*Generic Text interface*)<sup>69</sup> a la traducción de los documentos emanados del Tribunal de Justicia de las Comunidades Europeas (incluyendo aquí, obviamente, el Tribunal de Primera Instancia)<sup>70</sup>. Se está trabajando para cubrir todo el proceso de producción y gestión de los documentos emanados de este Tribunal, tanto los internos (la jurisprudencia) como los externos (las diferentes piezas que se integran dentro del procedimiento judicial). Así, el sistema desarrollado proporciona ayuda para la redacción de los documentos, para la inclusión de citas a otras sentencias o normas jurídicas, para la redacción de los fundamentos jurídicos,

---

<sup>64</sup> Es destacable el informe elaborado en abril de 2001 por esta agencia gubernamental sobre la importancia del estándar XML y la necesidad de las administraciones públicas del estado francés de desarrollar un repertorio de esquemas de acceso público con el doble objetivo de favorecer la interoperabilidad e intercambio electrónico de datos y facilitar la conservación de los documentos electrónicos. Véase en *De XML au répertoire de schémas XML de l'administration* [documento HTML]. Paris: ATICA, 12 avril 2001. Disponible en <http://www.atica.pm.gouv.fr/dossiers/documents/lettreXMLrepertoire.shtml> (consultado el 11 de abril de 2002).

<sup>65</sup> La información oficial relativa a estos proyectos de investigación aplicada en el desarrollo de esquemas y DTD XML se encuentra en <http://www.atica.pm.gouv.fr/XML/repertoire.shtml>

<sup>66</sup> <http://www.conseil-constitutionnel.fr/>

<sup>67</sup> La información oficial sobre este proyecto se encuentra disponible en la dirección [http://www.atica.pm.gouv.fr/XML/repertoire\\_de\\_schema/spad.shtml](http://www.atica.pm.gouv.fr/XML/repertoire_de_schema/spad.shtml)

<sup>68</sup> El documento electrónico en formato de MS-Word en el cual se encuentra transcrita esta DTD, además de la representación gráfica del modelo de contenido de cada uno de los elementos integrantes de la misma, está disponible en [http://www.atica.pm.gouv.fr/XML/documents/spad\\_docum.doc](http://www.atica.pm.gouv.fr/XML/documents/spad_docum.doc)

<sup>69</sup> Natalie Maurice, Directora de los servicios de traducción del Tribunal de Justicia de las Comunidades Europeas, presentó en el seno de la *Conferencia para una Infraestructura Terminológica en Europa* (París 13-15 de marzo de 2000) una interesante comunicación en la cual se explicaba con detalle el desarrollo y puesta en marcha de este proyecto. Natalie Maurice. "Intégration de ressources documentaires et linguistiques dans un processus de production multilingue de documents. Le projet GTi (Generic Text interface)" [documento HTML]. En: *Conférence pour une Infrastructure Terminologique en Europe* (2000. París). Disponible en <http://www.unilat.org/dtil/etis/actasTDCnet/maurice.htm> (consultado el 11 de abril de 2002).

<sup>70</sup> Toda la información oficial relativa al Tribunal de Justicia y al Tribunal de Justicia de Primera Instancia de las Comunidades Europeas se encuentra disponible en <http://www.curia.eu.int/>

para la traducción de todos estos documentos y, finalmente, para la comunicación y publicación de los mismos en los diferentes medios y espacios de difusión electrónicos e impresos. Todo este complejo entramado tecnológico se sustenta bajo una arquitectura de trabajo cliente-servidor donde XML se emplea como formato para el marcado, el almacenamiento y el intercambio de estos documentos.

En otros países del mundo, y para no extendernos demasiado en este apartado, se puede destacar el caso de Canadá, y de forma más concreta el Gobierno autónomo de Québec. Desde 1997 se han venido lanzando diversas iniciativas y planes estratégicos para el desarrollo tecnológico de sus administraciones públicas, caso del denominado *Chantier en ingénierie documentaire*<sup>71</sup>, proyecto llevado a cabo de forma conjunta por diferentes ministerios y organismos del gobierno de Québec (la Biblioteca Nacional, los archivos nacionales y el Grupo de responsables de gestión documental de dicho gobierno, principalmente) entre los años 1997 a 1999 con el objetivo principal de establecer una serie de directrices generales para la gestión integral del conjunto de documentos impresos y electrónicos gubernamentales. Es de destacar, por la importancia que tiene para nuestra disciplina, que en todos estos proyectos ha participado activamente el Grupo Departamental de Investigación en Estructuras Documentales de la Escuela de Biblioteconomía y Ciencias de la Información de la Universidad de Montreal, dirigido por el profesor Yves Marcoux<sup>72</sup>.

## Experiencias del grupo de trabajo

En el caso español, el colectivo de profesores adscritos al *Grupo de Tecnologías de la Información* del Departamento de Biblioteconomía y Documentación<sup>73</sup> de la Universidad Carlos III de Madrid, encabezado por la Dra. Mercedes Caridad Sebastián, viene desarrollando desde hace ya varios años una línea de investigación aplicada en el tratamiento y difusión de documentos jurídicos con las tecnologías propias de la Web.

Este grupo siempre ha confiado en las enormes posibilidades que estos lenguajes de marcado ofrecen para el tratamiento de los textos de carácter jurídico, lo que le ha llevado a poner en marcha en estos últimos años diversos proyectos de investigación aplicada al tratamiento y difusión de la documentación jurídica, en especial los documentos legislativos y jurisprudenciales, haciendo uso para ello de las mencionadas tecnologías. Ejemplos significativos de esta línea de investigación han sido los casos del proyecto desarrollado en 1998 (de diciembre de 1997 a mayo de 1998, exactamente), como contrato para actividades de asistencia técnica (art. 11 de la LRU) entre Dromi San Martino Consultores S.A. (del mismo grupo empresarial que la editorial Ciudad Argentina radicada en Buenos Aires, Argentina, receptora del producto final) y el Departamento de Biblioteconomía y Documentación de la Universidad Carlos III de Madrid, para la conversión de toda la normativa del Mercosur a un formato electrónico e hipertextual, capaz de ser difundido a través de Internet<sup>74</sup> y en soporte óptico (CD-ROM), así como el proyecto de investigación financiado por la Comunidad de Madrid en 1999 y desarrollado en 2000 para la creación de

---

<sup>71</sup> La información al respecto de este desarrollo se encuentra disponible dentro del sitio web de la *Autoroute de l'information* del gobierno de Québec en <http://www.autoroute.gouv.qc.ca/dossiers/ingenierie.htm>

<sup>72</sup> <http://mapageweb.umontreal.ca/marcoux/grds/>

<sup>73</sup> <http://www.bib.uc3m.es/>

<sup>74</sup> La versión electrónica se encuentra disponible en la dirección <http://hipatia.uc3m.es/~mercosur/>



una base de datos hipertextual de las Disposiciones Generales publicadas en el Boletín Oficial de esta comunidad autónoma<sup>75</sup>.

En fechas recientes (febrero de 2002) se ha defendido en esta Universidad una tesis doctoral realizada por el profesor Bonifacio Martín Galán y dirigida por el profesor Dr. J. Tomás Nogales con el título de *Tratamiento y Difusión en Internet de Información Jurisprudencial Mediante Tecnologías XML: Aplicación al Caso del Tribunal Constitucional*. Esta tesis doctoral ha desarrollado una maqueta inicial para el tratamiento de las Sentencias dictadas por el citado Tribunal mediante el uso de diversas tecnologías XML, construyendo de este modo un modelo de base de datos jurisprudencial disponible en la WWW alternativo a los modelos convencionales que se han venido desarrollando en las pasadas décadas<sup>76</sup>. Para ello, se han seguido diversos procesos como, entre otros, la definición formal de la estructura lógica de este tipo especial de documentos jurídicos, el marcado directo de las Sentencias, la creación de hojas de estilo para su presentación a través del lenguaje CSS (*Cascading StyleSheet Language*), la elaboración de mecanismos para la transformación de estos documentos en otros productos secundarios necesarios para la navegación y consulta de la información haciendo uso para ello de la tecnología XSLT (*Extensible Stylesheet Language-Transformation*), un acercamiento a la tecnología XLink (*Extensible Linking Language*) para el establecimiento de relaciones hipertextuales avanzadas o a la tecnología RDF (*Resource Description Framework*), la incorporación de un potente motor de búsqueda documental que produzca resultados precisos, así como la instalación y configuración de un servidor web especial (basado en la tecnología Cocoon del consorcio Apache<sup>77</sup>) capaz de procesar y difundir estos documentos en el espacio de publicación de la WWW.

Finalmente destacar que en estos momentos se está desarrollando dentro de este Departamento otra tesis doctoral llevada a cabo por la profesora M<sup>a</sup> del Carmen Arellano Pardo y dirigida igualmente por el Dr. J. Tomás Nogales, bajo el título de *La organización hipertextual del ordenamiento jurídico por medio de tecnologías XML*, cuya finalización está prevista para verano de 2002. En ella se aborda igualmente el uso de las tecnologías XML aplicadas en este caso a los documentos legislativos, incidiendo de manera especial en las aportaciones que el empleo del lenguaje extensible de hipertexto (XLink) hace al tratamiento y difusión de este tipo de documentos jurídicos.

En marzo de 2002 este equipo de investigación, junto con profesores del Departamento de Ingeniería Telemática de la misma universidad, ha iniciado una colaboración con el CENDOJ (Centro de Información Judicial) del CGPJ (Consejo General del Poder Judicial) para realizar un estudio preliminar sobre las posibilidades de aplicación de la familia de estándares XML a la creación de una base de datos que recogerá todas las resoluciones emanadas de los diversos órganos judiciales (Tribunal Supremo, Tribunales Superiores, Audiencias Provinciales...) y que se pondrá a disposición de sus potenciales usuarios a través de Internet, en el marco del proceso de modernización de la Justicia en España y en la Unión Europea en general.

---

<sup>75</sup> La versión electrónica se encuentra disponible en la dirección <http://hipatia.uc3m.es/~bocm/>

<sup>76</sup> Esta maqueta se encuentra disponible en Internet para su consulta en la dirección <http://damocles.uc3m.es/TC/>

<sup>77</sup> Más información sobre esta organización sin ánimo de lucro en <http://xml.apache.org/>

## **Conclusiones**

El uso de la tecnología Web/XML para la gestión y difusión de documentos jurídicos es muy conveniente, ya que alcanza una difusión planetaria a un coste mínimo, simplifica el acceso a los documentos relacionados al permitir la inserción de enlaces en las propias referencias, permite integrar todo tipo de medios en documentos complejos, se adapta mejor a estos documentos que las bases de datos convencionales y es tecnología normalizada, gratuita, muy soportada e independiente de fabricantes. Con lo expuesto, y en particular con los proyectos descritos, tanto los ya realizados como los que están en elaboración, creemos haber contribuido a demostrar la utilidad, e incluso la necesidad, de la aplicación de las tecnologías XML para el tratamiento y la difusión de información jurídica a través de Internet, promoviendo la puesta en marcha de servicios de información electrónica de altas prestaciones, que resulten útiles y manejables para el profesional o el estudioso del Derecho y, en definitiva, para el ciudadano, democratizando así el acceso a la documentación jurídica.