

Lenguajes de marcado y documentación jurídica: experiencias del grupo TecnoDoc de la UC3M

M.^a del Carmen Arellano Pardo

J. Tomás Nogales Flores

Universidad Carlos III de Madrid (España)

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 estas 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 haciendo uso de lenguajes de marcado como HTML y XML, 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 en I + D con empresas, proyectos de investigación con financiación pública y tesis doctorales. Estas experiencias han venido demostrando la importancia creciente de estas tecnologías propias de la World Wide Web para los profesionales de la documentación.

Palabras clave: Documentación jurídica. Informática jurídica documental. Documentos legislativos. Documentos jurisprudenciales. Lenguajes de marcado. HTML. XML. Internet. World Wide Web.

Abstract

Mark-up languages can be viewed as a complementary or alternative technology to conventional textual databases, but with remarkable benefits over these ones for the management of collections of large documents with complex and flexible structure and plenty of cross references, as is the case of legal texts. Experiences of processing and diffusion of legal information using HTML and XML are presented; they have been developed for several years by a working group in the Library and Information Science Department of Carlos III de Madrid University, taking the form of R&D consultancy contracts with companies, research projects with public funds and doctoral theses. These experiences have proved the increasing importance of these World Wide Web technologies for the information professionals.

Keywords: Legal information. Legal information systems. Legislative texts. Jurisprudential texts. Mark-up languages. HTML. XML. Internet. World Wide Web.

1. Presentación

En este trabajo se presentan algunos de los resultados de una de las líneas de investigación seguidas por el grupo de Tecnologías de la Información del Departamento de Biblioteconomía y Documentación de la Universidad Carlos III de Madrid entre los años 1998 y 2003. Cuando esta Universidad empezó a crear su catálogo de grupos de investigación reconocidos, el grupo citado, con algunos cambios en su composición, pasó a denominarse *Grupo de Tecnologías Aplicadas a la Información y la Documentación* (TecnoDOC). La línea de investigación mencionada tiene como objetivo fundamental el análisis de las posibilidades que conllevan los lenguajes de marcado descriptivo para el tratamiento de documentos (entendiendo por tales aquellos que contienen información predominantemente textual), y, en el caso concreto que ocupa a este trabajo, los de naturaleza jurídica, tanto textos legislativos como decisiones judiciales.

2. Contexto de la investigación

Las actividades desarrolladas dentro de esta línea de investigación se han centrado en distintas experiencias sobre la aplicación de los lenguajes de marcado descriptivo para el tratamiento, organización y recuperación de la información de distinta naturaleza, siendo destacable el caso de los textos literarios, periodísticos, archivísticos y jurídicos por haber sido tratados en el marco de proyectos de investigación financiados con fondos públicos o de contratos de I + D con empresas privadas. En algunos casos se ha tratado de aplicar lenguajes reconocidos como estándares internacionales, como TEI para las obras literarias, y en concreto para textos del teatro español del Siglo de Oro, o EAD para los instrumentos de descripción archivística, concretamente para un extenso catálogo de pergaminos. En el caso de la documentación jurídica, la propia evolución de los lenguajes de marcado descriptivo determinó que en las primeras experiencias el grupo utilizara HTML (HyperText Markup Language), mientras que posteriormente se aprovechó la capacidad de XML (Extensible Markup Language) para definir lenguajes de marcado específicos, al no existir en España estándares apropiados para esta tipología documental.

Los proyectos financiados, que se comentarán más adelante, han sido *Desarrollo de una versión electrónica del Código del Mercosur, Base de datos hipertextual, accesible en línea y en soporte óptico, de las disposiciones generales publicadas en el Boletín Oficial de la Comunidad de Madrid, Estudio sobre la utilización de XML para el archivo electrónico de las sentencias judiciales* (en el

CENDOJ) y, por último, *Desarrollo de una DTD de XML para el mercado de resoluciones judiciales (en el CENDOJ)*.

A estos proyectos habría que añadir dos tesis doctorales dirigidas por J. Tomás Nogales y realizadas por dos miembros del grupo TecnoDOC, ambos profesores del Departamento de Biblioteconomía y Documentación de la Universidad Carlos III de Madrid, y defendidas en esta misma Universidad: *Tratamiento y difusión en Internet de información jurisprudencial mediante tecnologías XML: aplicación al caso de Tribunal Constitucional*, realizada por Bonifacio Martín, y *La organización hipertextual del ordenamiento jurídico por medio de tecnologías XML*, elaborada por M.^a del Carmen Arellano.

En estas investigaciones siempre se tuvo presente que el hecho de que se analice la documentación jurídica solo desde el punto de vista de su tratamiento informático-documental no supone que sea el derecho el que deba adaptarse a las posibilidades que ofrecen los sistemas informáticos de organización de la información. En palabras de Aguiló Reglá (1988, p. 368),

no se trata de modificar el Derecho para satisfacer exigencias informático-documentales, sino de demostrar que, en realidad, se está ante problemas jurídicos. En otras palabras, conocer el alcance de una cláusula derogatoria no es un problema informático-documental, es un problema jurídico que, en la medida en que no pueda resolverse jurídicamente, supone disfunciones.

Las características propias de los textos jurídicos, tanto desde la óptica estrictamente jurídica como desde la documental, obligan a aplicar técnicas de análisis, tratamiento, almacenamiento y recuperación adaptadas a esas peculiaridades, lo que implica, en ocasiones, la adopción de soluciones poco convencionales para determinados problemas de tratamiento. Por ello, como paso previo a cualquier intento de creación y puesta en marcha de una base de datos documental integrada por este tipo de textos, se impone un análisis riguroso de sus características propias, de manera que sean las soluciones técnicas documentales e informáticas las que se adapten a los textos, y nunca a la inversa. Este principio de actuación ha regido las investigaciones llevadas a cabo por el equipo, tanto en sus primeros trabajos para la aplicación de HTML, el lenguaje genérico para los documentos difundidos vía web, como en los posteriores para la definición de lenguajes ad hoc basados en XML para tipos documentales específicos que no cuentan con estándares internacionales ni nacionales para su tratamiento, como es el caso de los textos legislativos y las sentencias judiciales.

A continuación se señalan algunas de las características de este tipo de documentos que de una forma u otra han tenido que ser consideradas para la definición y la elaboración de las soluciones técnicas que se han propuesto en las diversas investigaciones del grupo de trabajo.

3. Características de la documentación jurídica

El derecho, como norma que es para la expresión de hechos sociales de los cuales se derivan consecuencias jurídicas o actuaciones administrativas, ha tenido desde sus orígenes una importancia social, política y económica vital en todos los países y épocas. Debido a la conocida máxima latina *ignorantia legis (o iuris) non (o neminem) excusat*, imperante en los países que han mantenido la tradición de la *lex romana*, el derecho ha mantenido a lo largo de los siglos la obligación de ser conocido y, por tanto, difundido entre todos los ciudadanos afectados por las normas que regulan la práctica totalidad de los aspectos de su vida en relación con otros ciudadanos o con las administraciones de la comunidad a la que se adscriben (Nogales et ál., 2003a).

Esto implica que una de las características que diferencian a la norma legislativa de cualquier otro documento es la necesidad de su publicación para que sea considerada como tal norma. La publicación es el medio que garantiza una comunicación y, por lo tanto, posibilita el conocimiento (Rodríguez-Zapata, 1987).

La Constitución española de 1978 (a partir de ahora mencionada como CE) establece esta necesidad de publicación en los artículos 9.3, “La Constitución garantiza [...] la publicidad de las normas”; 91, “El Rey sancionará en el plazo de quince días las leyes [...] y ordenará su inmediata publicación”; 96.1, “Los tratados internacionales [...] una vez publicados oficialmente en España” y en su Disposición Final, “Esta Constitución entrará en vigor el mismo día de la publicación de su texto oficial en el Boletín Oficial del Estado. Se publicará también en las demás lenguas de España”. Cabe destacar que estas referencias a la publicación de las normas no establecen el medio en el que se debe realizar dicha publicación, salvo en el caso de la propia CE. Hasta la fecha, la forma elegida es el periódico oficial en papel, cuyo nombre ha sufrido varias modificaciones; las dos denominaciones más conocidas son la de *Gaceta de Madrid* (1836-1936) y la actual de *Boletín Oficial del Estado* (1936-). De esta forma, los legisladores dejan abierta la posibilidad de que la publicación pueda llevarse a cabo en otro medio o con otro nombre (Sainz, 1993).

Una vez que los legisladores han cumplido su responsabilidad de dar publicidad a la norma por medio de su publicación, la responsabilidad del conocimiento de dicha norma recae sobre el ciudadano, según queda expresado en el artículo 6.1 del Código Civil en la traducción de la máxima latina: “La ignorancia de las leyes no excusa de su cumplimiento”. Esto implica que la información recogida en las normas es de interés general, puesto que afecta a todo tipo de personas y organizaciones públicas o privadas sin tener en cuenta particularidades económicas, sociales, políticas o de otro tipo. Siendo así, quizá no es necesario u obligado por la CE pero sí recomendable que las normas no solo sean publicadas por el Estado en los cauces, medios y soportes tradicionales, sino que además lo sean de la forma más accesible para el conjunto de la ciudadanía. Los nuevos medios y soportes, y

muy especialmente Internet, pueden ser aprovechados a tal fin de manera habitual, como en su día, de forma excepcional, se difundió en los medios de comunicación la primera convocatoria democrática.

Con respecto a las características formales de los documentos jurídicos nos encontramos con textos de extensión muy variable, que pueden oscilar entre unas pocas líneas y cientos de páginas. Por otro lado, tales documentos pueden incluir, junto a lo que propiamente podemos considerar como texto, tablas (que aunque contengan datos textuales o numéricos tienen una disposición especial que es significativa), gráficos o imágenes (especialmente los normativos existentes en reglamentos o especificaciones). Los lenguajes de marcado y los sistemas de gestión de bases de datos que explotan esta nueva forma de tratamiento documental aportan soluciones idóneas para la gestión adecuada de todas las características señaladas (Nogales et ál., 2000a).

De especial importancia es la existencia dentro de estos documentos de múltiples referencias a otras partes del mismo documento, a otros documentos o a partes concretas de otros documentos (Hernández, 1998). En el caso de las resoluciones judiciales, además de incluir referencias a las normas en que se sustentan, pueden remitir (sobre todo en el caso de ciertos países) a sentencias previas de otros órganos judiciales. La tecnología de enlaces hipertextuales integrada en los lenguajes de marcado, en especial la más avanzada que se ha generado en el ámbito de XML, da una solución a la expresión de estas referencias y a su explotación para ofrecer un acceso inmediato a los documentos referenciados.

Los textos legislativos, además, son susceptibles de modificación o derogación por otras disposiciones, por lo que podríamos decir que se trata de “documentos vivos”. En estos casos no solo disponemos de referencias expresas de un texto a otro que es modificado (o derogado), sino que además sería sumamente interesante incluir en este la referencia a la norma modificadora, junto con el propio texto modificado. De esta manera, como se detallará más adelante, un texto admite una lectura dependiente de una fecha concreta de vigencia. También en la tecnología de los lenguajes de marcado pueden encontrarse soluciones a esta problemática (Arellano et ál., 2001).

Según apunta Otto (1988, p. 80), el principio de interpretación de las normas “tiene su primer fundamento en el de unidad del ordenamiento jurídico. Dado que la Constitución fundamenta la validez de todo el ordenamiento jurídico, las normas que lo integran no pueden contradecirse entre sí, y si en realidad se produce esa contradicción el conocimiento jurídico está obligado a eliminarla”. Esta unidad se define por el conjunto de disposiciones coordinadas, de forma que el artículo de una norma puede carecer de sentido o resultar insuficiente si se aísla del contexto de la propia norma, de las complementarias o de las de general aplicación (López-Muñiz, 1984, p. 173). Esta coordinación de disposiciones se establece por

medio de referencias entre ellas, convirtiéndose en el mecanismo usado por el legislador para garantizar la unidad del ordenamiento. Por una parte, encontramos las relaciones que responden a la unidad estructural: de fundamento, antecedentes y competencia en virtud de la cual se dictan; de aplicación o remisiones externas, con las que se extiende el ámbito de competencia de una disposición ya existente a los casos recogidos en la nueva disposición; y de desarrollo, por las que se especifican los medios y los casos de aplicación de una regulación existente. Esta tipología de las relaciones puede quedar expresa en los textos informatizados y puede ser aprovechada en la recuperación gracias a las posibilidades que aportan los lenguajes de marcado.

Otra característica de los documentos jurídicos que se ha de tener en cuenta para su tratamiento, organización y recuperación es su estructura formal, es decir, la división ordenada de las materias que regulan mediante fórmulas epigráficas de separación que sirven de soporte contextual o instrumental a la sistematización lógica del mismo (Cazorla, 1999). De nuevo, los lenguajes de marcado se muestran como la mejor solución: en su propia naturaleza está la capacidad para la definición de estructuras de cualquier tipo, desde las más rígidas a las más flexibles, y de cualquier nivel de profundidad del análisis (Arellano, 2003).

En relación con la estructura de los documentos jurídicos, en cuanto a los documentos legislativos puede generalizarse a las normas con rango de ley lo establecido en la resolución de 15 de noviembre de 1991 por la que se dispone la publicación del acuerdo del Consejo de Ministros según el cual se aprueban las directrices sobre la forma y estructura de los anteproyectos de ley (BOE n.º 276 de 18 de noviembre de 1991). En dicho acuerdo no solo se especifican las cuatro divisiones principales, sino que se precisa tanto el epígrafe o nombre con el que deben identificarse como el tipo de contenido de cada una.

Según estas directrices, toda norma debe comenzar su redacción por su título formal, en el que además de exponerse de forma clara, completa y concisa el objeto de esta, deben incluirse los datos identificadores de la misma, es decir, el rango, el número oficial y la fecha de aprobación.

A continuación debe aparecer la exposición de motivos, donde se manifiestan las razones que justifican la incorporación de la norma al ordenamiento jurídico. Es frecuente que estas razones se fundamenten en referencias a normas que complementa y a las normas de general cumplimiento (Martín, 1989).

La tercera parte en la que se estructura una norma es el articulado, formado por todos los artículos en los que se expresa la regulación. Dependiendo de la extensión de los artículos, estos pueden a su vez dividirse en apartados, que además podrán tener otro nivel de división. Según el rango de la norma y el número de artículos, estos pueden organizarse de acuerdo a una estructura jerárquica cuyo nivel

superior sería el *título*, seguido por el *capítulo* y, por último, la *sección*; esta última es la única que puede aparecer o no; si no lo hace, los artículos quedan directamente agrupados en capítulos (Sánchez, 1989).

La norma debe acabar con la redacción de las disposiciones. En este caso el nombre formal no se refiere al conjunto de disposiciones, sino a cada uno de sus tipos: *adicionales*, *transitorias*, *derogatorias* y *finales*; es decir, el nombre formal sería para las primeras disposiciones adicionales, y así sucesivamente. La aparición de cada uno de los tipos de disposiciones no es necesaria, pero cuando figuran suelen hacerlo en este orden, e incluso puede existir más de una de un mismo tipo. En general, las adicionales contienen aspectos colaterales al tema fundamental que regula la norma, las transitorias detallan el régimen de aplicación con respecto a la normativa vigente hasta ese momento, las derogatorias señalan aquellas normas que son derogadas y las finales recogen la entrada en vigor y, en el caso de que no haya derogatorias, las normas que pierden su vigencia (Viver, 1989). Dado que el conjunto de estas disposiciones incluye parte de la regulación de una norma diferenciada de la regulación principal desarrollada en el articulado, en las propuestas de este equipo de investigación se las ha agrupado formando la última división de primer nivel dentro de la estructura de una norma (Nogales et ál., 2004).

En cuanto a los documentos judiciales, ciñendo el análisis a un tipo concreto, en las sentencias del Tribunal Constitucional (en adelante TC) se pueden distinguir, en un primer nivel de definición de la estructura, un total de seis elementos. En una primera parte, que denominamos *cabecera*, se sitúa aquella información que, siendo de interés para su posterior control y tratamiento, no viene directamente incluida en el texto de las sentencias. Dentro de la propia sentencia, que engloba los cinco elementos restantes, diferenciamos como *preámbulo* el segundo bloque, la parte en la que se introduce la sentencia, informando sobre la composición de la sala que ha juzgado el caso y dictado dicha sentencia, así como los datos característicos del recurso interpuesto ante el TC. A continuación aparecen los datos identificados como *antecedentes*, el tercer elemento, donde se señala la historia del acontecimiento judicial tratado hasta llegar a su revisión por parte del TC. La cuarta división de la estructura de la sentencia recoge los *fundamentos jurídicos*, donde se indican los principios doctrinales puestos de manifiesto y la normativa aplicada en la resolución. Una vez fundamentado, se expone el *fallo*, quinto elemento, donde el Tribunal dicta su fallo y hace su correspondiente disposición. Por último se identifican los *votos*, sexto elemento, opcional, pues no en todas las sentencias se producen votos particulares emitidos por algún magistrado, discrepando o matizando lo dispuesto por la sala o el pleno del TC (Nogales et ál., 2002).

Por último, volviendo a las características generales de los documentos jurídicos, el hecho de que puedan ser consultados con fines distintos y por diversos tipos de usuarios hace muy conveniente, y a veces incluso imprescindible, disponer

de diferentes formas de presentación de la información contenida en cada uno de ellos, ya sea sentencia o norma, o bien en un grupo, por ejemplo, de normativa que regula una materia.

En las resoluciones judiciales puede y, de hecho, suele aparecer información sensible, como datos de identificación o localización de personas físicas implicadas o citadas en el procedimiento judicial, ya sean partes en una demanda, apelación, denuncia o querrela, o bien víctimas, testigos, peritos o notarios; esta información debe ser sustituida u ocultada en todo el documento, independientemente del apartado en el que se encuentre, mediante un proceso de “vaciado”, a efectos de que su difusión pública no lesione los derechos de los afectados. No se sometería a este proceso la información referida a los agentes judiciales (jueces, magistrados, abogados, procuradores, etcétera) que intervienen, salvo cuando aparecen como implicados en el caso, ni los datos de identificación de personas jurídicas. Entre los datos sensibles pueden señalarse los nombres y apellidos, presentados de forma conjunta o independiente, o cualquier otra forma de identificación nominal como alias, motes o apodos; los números de identificación o códigos de pertenencia a cualquier colectivo (DNI, NIF, CIF, número de afiliación a la Seguridad Social o de pasaporte, números de agentes, colegiados, etcétera); las localizaciones geográficas de ámbito inferior al de una población, como nombres de calles, números de bloques, portales o pisos; códigos de identificación de carreteras y puntos kilométricos; números o nombres de identificación de fincas; y los códigos de identificación concreta de vehículos, como matrículas, números de bastidor, nombres de barcos, pero no los datos de identificación genérica, como marca y modelo. Los lenguajes de marcado, una vez más, permiten incluir mecanismos para la sustitución u ocultación de ciertos datos en una opción de representación de un documento, manteniendo la información original o completa en otra opción de representación, disponible, por ejemplo, para un grupo determinado de usuarios. Un mismo documento puede ser representado en un número ilimitado de formas (Nogales et ál., 2003b).

En el caso de la legislación, el jurista puede estar interesado en la consulta de la normativa vigente en la actualidad, o bien en la de aquellas disposiciones de aplicación en una fecha determinada, más que en la consulta de los textos normativos tal cual fueron publicados en su día en el boletín o diario oficial que le correspondiera. Otro caso de interés histórico podría ser el interés por conocer la evolución sufrida en la regulación de una misma materia a lo largo del tiempo (López-Muñiz, 1984). Cualquiera de ellos podría suponer la consulta de la legislación con una presentación diferente a la que tuvo en origen, compuesta de distintas redacciones publicadas en distintas normas. Con los lenguajes de marcado pueden utilizarse los mecanismos oportunos para obtener una visión sincrónica o diacrónica de una norma que ha sufrido modificaciones en el tiempo (Arellano, 2005).

4. Experiencias con legislación

A continuación se presentan las distintas experiencias de tratamiento de información jurídica llevadas a cabo por el grupo de investigación. Cada una de ellas aborda una problemática particular y aporta unas soluciones específicas, derivadas de sus propios objetivos, de los intereses de terceros en el caso de contratos de investigación, e incluso del estado de la propia tecnología en el momento en que se desarrolla. En común estas investigaciones responden al método deductivo experimental con validación práctica.

4.1. El Código del Mercosur

Este proyecto se lleva a cabo como contrato de investigación entre el grupo Dromi Sanmartino Consultores – Ediciones Ciudad Argentina y el Departamento de Biblioteconomía y Documentación de la Universidad Carlos III de Madrid, entre enero y junio de 1998. La finalidad general era desarrollar una versión electrónica del Código del Mercosur en forma de base de datos con organización hipertextual susceptible de ser difundida tanto en soporte óptico como a través de la Web.

El objeto de la investigación era la compilación de textos emanados del Mercosur en el periodo comprendido entre 1991, año de su creación, y 1998, fecha en la que se llevó a cabo el proyecto. Tal compilación existía hasta ese momento solo en soporte papel, e integraba una colección de más de novecientos textos jurídicos de derecho comunitario entre protocolos y acuerdos, declaraciones presidenciales, decisiones del Consejo del Mercado Común, resoluciones del Grupo del Mercado Común, el bloque más numeroso, y directivas de la Comisión de Comercio del Mercosur. A esto se añadía un estudio preliminar, así como varios índices (de materias, normativo, cronológico, bibliográfico y de abreviaturas), hasta completar un total de más de diez mil páginas impresas.

La primera fase de la metodología del proyecto supuso el análisis de la documentación, para determinar tanto su tipología general como la casuística particular. Dentro de la colección de textos a tratar se identificaron tres tipos documentales. Por una parte, el estudio preliminar, con una estructura propia de una monografía, con algunas referencias internas (a epígrafes concretos, notas explicativas, etcétera), y múltiples referencias a la normativa. El segundo de los tipos documentales, el fundamental en el proyecto, lo constituía el conjunto de normativa emanada del Mercosur, con las características propias de los textos legislativos ya mencionadas en el apartado anterior, de las que solo cabe destacar en este caso la ausencia de referencias externas al corpus documental a tratar, dado que se trabajaba con la totalidad de la normativa emanada hasta ese momento y que el Mercosur no tiene competencias judiciales. Por último, los índices, que por su propia naturaleza contenían miles de referencias a la normativa; además, en el caso del de materias, también incluía multitud de referencias internas por la remisión de unas voces a otras.

A continuación se analizaron los archivos originales generados con el *software* PageMaker, programa de maquetación de páginas profesional. Dada la aplicación empleada, los archivos mostraban un formato exclusivamente orientado a la impresión (cajas de texto diferenciadas para los distintos estilos o visualizaciones del texto, etcétera), y además no existía ninguna correspondencia entre los distintos archivos físicos, simples grupos de páginas, y los documentos lógicos, por ejemplo cada uno de los textos legislativos.

Se estableció una nomenclatura autoexplicativa y válida para cualquier plataforma, tanto para la identificación del archivo como para la estructura lógica del corpus documental. Con ello se consiguió una localización de los textos normativos dentro del sistema de almacenamiento que posibilitaba que las rutas de acceso se definieran a partir de los datos identificativos de la propia norma, tipo de norma y número oficial (la numeración oficial en los documentos legislativos es correlativa dentro de un mismo año). Por ejemplo, la Resolución 1/1994 tenía la ruta de acceso */normativa/r/1994/r94001.htm*, donde se aprecia que tanto la ubicación en el sistema de archivos como el nombre del propio archivo identifican y denominan a cada norma. Esto supuso una enorme ventaja a la hora de establecer y expresar las rutas de los enlaces a normas citadas en un texto concreto que se estuviera tratando en un momento dado, aun cuando tales normas citadas no hubieran sido incorporadas todavía a la base de datos, ya que los datos de la referencia permitían identificar la localización de archivos no existentes todavía.

El tratamiento de la información comenzó con la conversión semiautomática de los formatos de archivo desde PageMaker a HTML directamente, o bien desde PageMaker a Microsoft Word de forma automática y la conversión a texto plano para su marcado en HTML. Además de estas conversiones se dan también, en esta fase del proyecto, las uniones o divisiones necesarias para conseguir una correspondencia entre archivos físicos y documentos lógicos hasta dejar, generalmente, uno solo por texto legal, con la salvedad de que, pensando en las limitaciones de velocidad de acceso a la red o de visualización de algunos navegadores de la época, se hizo conveniente la división de documentos lógicos extensos, de algunos cientos de páginas, en varios más pequeños, en cuyo caso se procedía a la creación de un índice interno en todos ellos con enlaces a los restantes. Fue necesario un tratamiento no automático de los documentos para la redefinición de tablas y el marcado de la estructura (los epígrafes internos de cada documento) y para la extracción, identificación y asociación de imágenes, así como la modificación de la resolución de algunas imágenes o su digitalización desde la edición del Código del Mercosur en su versión original en papel. Se desarrollaron pequeños programas de *scripts* para la automatización del marcado básico de la estructura común de los documentos. Sin embargo, la inserción de los enlaces, tanto internos como externos a cada documento, en número de varias decenas de miles, presentes tanto en los in-

dices como especialmente en la normativa, hubo que realizarla de forma manual, dadas las dificultades insalvables que se derivaban de la ausencia de normalización en las referencias del texto.

El producto final del proyecto tuvo un diseño basado en dos *frames* o marcos, de los cuales el superior, siempre visible, facilitaba de forma permanente el acceso a las distintas partes de la obra, así como a cada uno de los índices, mientras que el inferior quedaba reservado para la presentación de los documentos. En el producto que comercializó Dromi Sanmartino se modificaron algunos aspectos del diseño, como la conversión del título de cada norma de texto a imagen, la inserción de un tercer marco lateral con el acceso a las partes de la obra, la eliminación del color de resalte en los avisos de derogación o modificación, la reducción del espacio de visualización de los documentos y, como consecuencia necesaria, de la resolución de las imágenes, entre otras. También se introdujeron cambios en la estructura de almacenamiento, con cambios en la nomenclatura de algunos directorios. Todas estas modificaciones, que en principio iban a ser meramente estilísticas, implicaron, a juicio del equipo investigador, una merma apreciable de la utilidad del producto final del proyecto.

El proyecto supuso la primera experiencia en la aplicación del marcado de texto descriptivo a los documentos jurídicos. Los lenguajes empleados fueron HTML para el marcado estructural y semántico y CSS (Cascading StyleSheet Language) para las características de representación, dado que en el momento de realización del proyecto XML estaba finalizando su desarrollo y los navegadores aún necesitarían unos años para trabajar con esta tecnología. Sin embargo, como lenguaje de marcado, incluso HTML, a pesar de ser tan genérico, demostró su capacidad y sus virtudes para el tratamiento con este tipo de documentos. Permitía la integración de los enlaces de las referencias en el propio texto, tanto las internas al documento (explícitas en el original) como las externas a documentos anteriores (también explícitas en el original) o las externas a documentos posteriores (derogaciones o modificaciones que eran incorporadas al original). Además, hacía posible tratar con documentos de cualquier tamaño, así como integrar distintos medios como texto, tablas o imágenes. Y, en cuanto a la publicación de los documentos tratados, permitía la difusión del producto tanto en soportes ópticos o magnéticos como a través de Internet. Por último, se valoró positivamente la posibilidad de integrar con cierta facilidad motores de búsqueda para la recuperación por texto completo, cubriendo así las deficiencias que pudieran encontrarse desde el acceso por el índice de materias (Nogales et ál., 2000b).

4.2. Disposiciones generales del BOCM

Este segundo proyecto de tratamiento de la documentación legislativa se llevó a cabo en el marco de una convocatoria pública de la Comunidad de Madrid en

2000, con objeto de desarrollar una base de datos hipertextual, accesible en línea y en soporte óptico, de las disposiciones generales publicadas en el *Boletín Oficial de la Comunidad de Madrid* (BOCM). El corpus documental a tratar estaba compuesto por 458 disposiciones.

La metodología comenzaba, como en el proyecto anterior, con un análisis de la tipología documental objeto de la investigación. En este proyecto la tipología quedaba identificada por los rangos normativos, ya que todos los documentos a tratar eran textos legislativos: leyes, decretos, resoluciones, órdenes y acuerdos. Si bien se pueden observar algunas diferencias entre los distintos rangos, como que el número oficial en leyes y decretos está normalizada por número correlativo dentro de un mismo año, mientras que para resoluciones y órdenes su aparición no está garantizada, también lo es que existen muchas coincidencias obvias, como el alto grado de variación en cuanto a su extensión y la existencia de múltiples referencias a otras normas.

En este proyecto también se disponía de una versión en formato digital de los documentos a tratar. En algunos casos se contaba con archivos HTML, mientras que en otros la fuente de datos estaba disponible en formato PDF. En ambas opciones era necesario el tratamiento de los archivos para su recomposición, puesto que ninguna de ellas nos ofrecía una correspondencia entre archivo físico y documento lógico, ya que cada archivo se correspondía con una página del BOCM impreso.

Se estableció una nomenclatura para los ficheros y para la estructura de almacenamiento basada en los mismos principios del proyecto anterior, esto es, que fuera autoexplicativa y válida para cualquier plataforma informática y sistema de archivos. Se creó un primer nivel de almacenamiento para los distintos rangos normativos, aunque en el caso de resoluciones y órdenes se estableció un segundo nivel por órgano emisor, y finalmente, en un último nivel, se agruparon por años para todos los rangos.

La fase de conversión de archivos físicos a documentos lógicos se realizó con el mismo grado de automatización señalado para el proyecto anterior, con la salvedad de que, en razón de la extensión de los documentos, no fue necesario dividir en distintas partes ninguno de ellos. El hecho de contar con algunos archivos en formato HTML no supuso grandes variaciones en la fase de marcado de la estructura de la norma, puesto que, al ser uno por página impresa, en ningún caso recogían la estructura global de la norma. Se elaboró un *script* para la inserción del marcado básico con elementos comunes al conjunto del corpus documental para su presentación tanto de forma visible en el navegador como para proceso informático, en forma de metadatos. La asignación de la estructura de cada documento, así como la definición de enlaces en las referencias tanto internas como externas, se realizó manualmente, al igual que en el proyecto anterior. La novedad en esta fase de la metodología consistió en el uso extensivo de clases, introducidas por

medio del atributo *class* a casi todos los elementos HTML, con la doble finalidad de, por un lado, precisar las características de presentación mediante hojas de estilo CSS y, por otro, posibilitar la conversión automática del marcado HTML a un lenguaje cualquiera derivado de XML y definido en una DTD (Document Type Definition) propia. Se trabajó con la premisa de que el marcado XML se impondría en lugar del HTML para colecciones de documentos de tipos y estructuras bien definidos, puesto que su naturaleza de metalenguaje y por ende su capacidad para definir lenguajes específicos permite reflejar las particularidades propias de cualquier tipo dado de documentos. Por otro lado, además de suponer una ventaja en la definición de la estructura y la semántica, también ofrece la posibilidad de desarrollar un sistema hipertextual más potente con la inclusión de enlaces bidireccionales, a múltiples destinos y de usuario, definidos con XLink (XML Linking Language) y XPointer (XML Pointer Language).

El producto final resultante siguió los mismos criterios de diseño, centrados en una presentación clara de la información y un acceso fácil y directo a las herramientas diseñadas para la localización de la normativa, índices cronológicos de emisión y publicación, por rango y por organismo emisor, siempre disponibles para el usuario en la parte superior de la pantalla, mientras que la inferior se reservó para la consulta del documento (Nogales, 2000c).

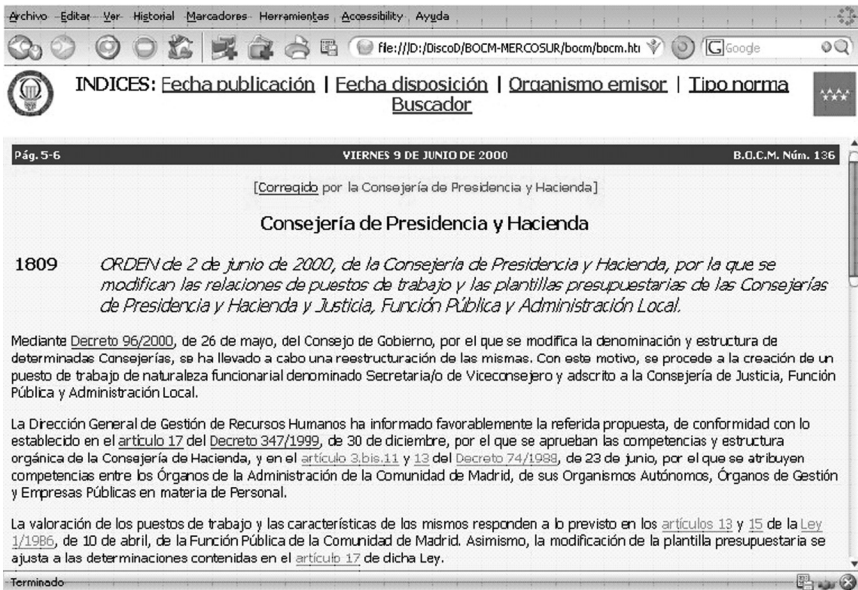


Fig. 1. Legislación del BOCM. Un ejemplo de acceso por índices.

4.3. Normativa del IRPF

Esta investigación se desarrolló en el marco de la tesis realizada por M.^a del Carmen Arellano y defendida en abril de 2003 bajo el título *La organización hipertextual del ordenamiento jurídico por medio de tecnologías XML*. Esta tesis tenía por objeto realizar el diseño de un modelo de hiperdocumento que, aplicando las tecnologías XML, fuera capaz de dar respuesta a las necesidades que plantea un sistema de información jurídica documental aprovechando las posibilidades que ofrece el hipertexto.

Este modelo fue aplicado en concreto a un conjunto de normas relacionadas que regulaban el Impuesto de la Renta de las Personas Físicas (IRPF) desde 1999 hasta 2001, normas que contenían la redacción vigente hasta enero de 2002. La selección del corpus documental se realizó teniendo en cuenta que las normas escogidas para formar parte de la maqueta presentada en la tesis debían tener un alto grado de relación, que además estas relaciones habían de recoger las distintas tipologías sobre las que se basa la unidad temporal y sistemática del ordenamiento jurídico, que tenía que darse al menos un caso de más de dos redacciones para un mismo precepto, que al menos una de las normas debía regular una materia de manera lo suficientemente extensa como para que su sistematización se realizase en varios niveles jerárquicos, y que los archivos de estas normas habían de estar ya disponibles en formato electrónico, ser de libre acceso y, si era posible, tener un formato que ayudara a su marcado en XML.

La nomenclatura empleada para la identificación de archivos y directorios seguía el criterio establecido desde la primera experiencia en el marcado de documentos legislativos, en el sentido de ser autoexplicativa y multiplataforma. La variación que se introdujo fue la permutación en el orden de aparición entre el número y el año de la norma para facilitar su localización dentro del directorio al agrupar todas las normas de un mismo año y ordenarlas por su número. Visto a través de un ejemplo, la Ley 40/1998 se llamará *l-1998-0040.xml* y se localizará en la ruta */normas/l/l-1998-0040.xml*. Puede apreciarse que se ha expresado el número oficial con cuatro dígitos, suficientes para reflejar todas las normas de un mismo rango emitidas en un año concreto.

En este caso se partía de archivos HTML en los que sí se daba una correspondencia con los documentos lógicos. Además, la conversión al marcado XML se veía facilitada por la posibilidad de incluir una cierta automatización, ya que el código HTML de partida identificaba los niveles estructurales de *título* y *capítulo* por medio de atributos *class*. El resto de las etiquetas HTML no podían ser reemplazadas por los elementos de la DTD creada ad hoc en el marco de la tesis, pues no identificaban los artículos y las listas tampoco identificaban a sus niveles inferiores de forma que pudieran ser reemplazadas por los elementos *apartado* o *subapartado*. Una vez que se habían introducido los elementos *título* y *capítulo* se

eliminaba el código HTML. La identificación del elemento *artículo*, y por tanto la inserción automática de las oportunas etiquetas de marcado (la final, </Artículo>, y la inicial, <Artículo>), también era posible gracias a que este elemento estructural siempre comienza con la palabra *Artículo* y solo los epígrafes de los artículos contienen esta palabra con la primera letra en mayúsculas. La etiqueta de cierre anterior al primer artículo y la de inicio posterior al último se eliminaban manualmente. El resto de los elementos y atributos XML también se introducían de modo manual.

La definición de los elementos de enlace para las referencias entre normas se realizaba también con una cierta automatización por medio de un *script*. En él se identificaban las citas a otras normas gracias al número oficial y al rango; se detectaban cadenas de caracteres con una cantidad de dígitos variable, seguidos de una barra (/) y cuatro dígitos finales, y una vez localizada una de estas cadenas se comprobaba que los caracteres anteriores se correspondieran con alguno de los rangos normativos, de manera que no solo se identificaba el rango que es parte del nombre del archivo que almacena la norma citada, sino que también se comprobaba que era una cita a una norma y no, por ejemplo, un intervalo de años; a continuación se insertaban las etiquetas propias de los elementos de enlace con los datos relativos a la localización de la norma, mientras que los relativos a la identificación del tipo de relación jurídica que hubiera motivado la cita se insertaban manualmente.

```

<Referencia xlink:type="extended" Tipo="modificacion">
  <TextoRef
    xlink:type="locator"
    xlink:href="norma1.xml#art1"
    xlink:label="texto"
    xlink:title="nueva redacción del artículo 7
      de la norma 2"/>

  <ObjetoRef
    xlink:type="locator"
    xlink:href="norma2.xml#art7"
    xlink:label="objeto"
    xlink:title="artículo 7 de la norma 2"/>

  <Desplazamiento
    xlink:type="arc"
    xlink:from="texto"
    xlink:to="objeto"/>

  <Desplazamiento
    xlink:type="arc"
    xlink:from="objeto"
    xlink:to="texto"/>

</Referencia>

```

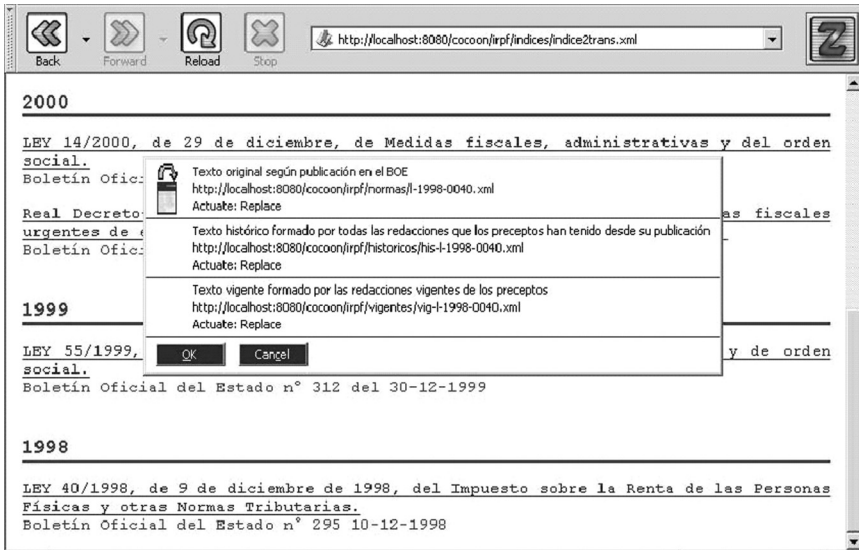



Fig. 2. Legislación del IRPF. Posibilidades del enlace definido con XLink.

Dada la imposibilidad de que el destino de un enlace se mostrase embebido dentro del documento que contuviera dicho enlace, era necesario usar distintos archivos XSLT (Extensible Stylesheet Language-Transformation) con el fin de definir las diferentes visualizaciones que pudieran ser interesantes de consultar sobre la regulación de una materia, vigencia en un momento dado, texto publicado o recopilación de las distintas redacciones de una materia a lo largo de la historia. Para la generación semiautomática de los diversos índices de acceso a la normativa (fecha de aprobación, fecha de publicación y rango) también fue necesario el uso de XSLT.

El producto resultante se basa totalmente en las tecnologías XML, algunas de las cuales no estaban implementadas en los navegadores de uso más extendido (Internet Explorer, FireFox u Opera) en el momento del cierre de la investigación; por ejemplo, la transformación con XSLT y otras siguen sin estarlo en la actualidad, como los enlaces complejos de XLink. Por ello fue necesario publicar el resultado en un servidor al que se le había instalado el marco de trabajo para XML desarrollado por el consorcio Apache, Cocoon, que realizaba las transformaciones de XSLT en servidor, y además acceder al sistema por medio del navegador DocZilla, que sí implementa casi en su totalidad las opciones de enlaces complejos de XLink y los punteros definidos por XPointer (Arellano et ál., 2003).

5. Experiencias con jurisprudencia

5.1. Sentencias del Tribunal Constitucional

En febrero de 2002 se defendió la tesis doctoral realizada por Bonifacio Martín Galán con el título *Tratamiento y difusión en Internet de información jurisprudencial mediante tecnologías XML: aplicación al caso del Tribunal Constitucional*. En el marco de esta tesis se ha desarrollado una maqueta inicial para el tratamiento de las sentencias de dicho Tribunal mediante el uso de tecnologías XML. El corpus documental de la maqueta está compuesto por siete sentencias emitidas entre los años 1982 y 2000.

Entre las fases principales para la elaboración de la experiencia pueden señalarse la definición formal de la estructura lógica de este tipo concreto de documentos jurídicos y su posterior plasmación en una DTD de XML, el marcado manual de las sentencias y la creación de hojas de estilo para su presentación a través del lenguaje CSS, así como la elaboración de mecanismos basados en XSLT para transformar estos documentos en productos secundarios necesarios para la navegación y consulta de la información (como índices diversos y resúmenes de las sentencias), además de las más convencionales para convertir a HTML. También se proponía un acercamiento a XLink para el establecimiento de relaciones hipertextuales avanzadas o a RDF (Resource Description Framework) para facilitar la incorporación de un motor de búsqueda documental que produzca resultados más precisos. La plataforma elegida para el sistema incluía la instalación y configuración de un servidor web, basado también en la tecnología Cocoon del consorcio Apache, capaz de procesar y difundir estos documentos en el espacio de publicación de la Web.

El empleo de XSLT fue fundamental para la generación automática de otros documentos XML secundarios o complementarios al sitio web desarrollado y, así, partiendo de los documentos XML marcados y haciendo uso de diversos ficheros XSLT se generaba automáticamente un documento abreviado de cada sentencia con sus datos más característicos y, por tanto, de consulta más sencilla y directa. También se empleó esta tecnología para la generación automática de diversos índices de acceso a las sentencias del Tribunal Constitucional, como el cronológico, por salas y por materias. De igual modo se podrían haber generado otros; por ejemplo, por tipo de recurso planteado o por fecha de publicación en el *Boletín Oficial del Estado*. De este modo, a medida que se van incorporando nuevas sentencias el sistema puede integrar de forma automática en los distintos índices la referencia a las mismas. Además se definieron distintos formatos de salida generados automáticamente a partir del texto marcado con XML a archivos HTML y PDF. Dicha conversión se consiguió por medio de XSL (Martín, 2002).

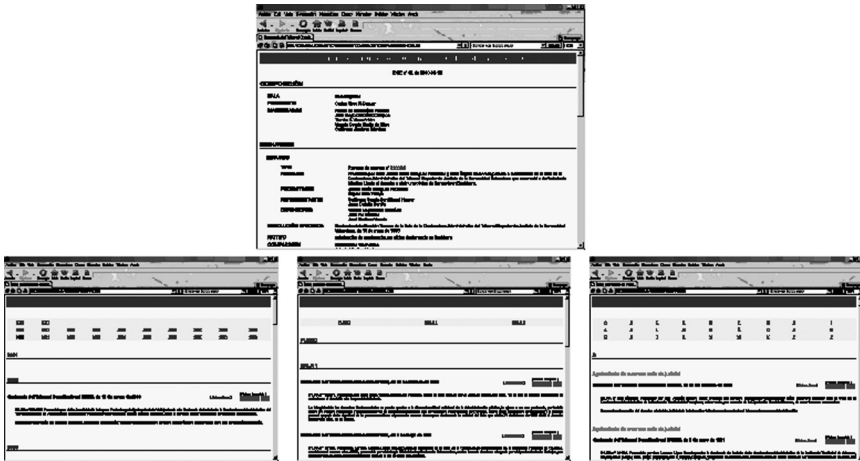


Fig. 3. Resoluciones del Tribunal Constitucional. Distintas formas de presentación gracias al marcado con XML.

5.2. Resoluciones judiciales tratadas en el Centro de Documentación Judicial (CENDOJ)

Esta experiencia se desarrolla en el marco de un contrato de asesoría entre la empresa PSd (Proyectos y Servicios Documentales) y los Departamentos de Ingeniería Telemática y Biblioteconomía y Documentación de la Universidad Carlos III de Madrid, cuyo destinatario final es el Centro de Documentación Judicial (CENDOJ) del Consejo General del Poder Judicial (CGPJ), para la aplicación de las tecnologías derivadas del marcado XML a las resoluciones judiciales emanadas de distintos órganos judiciales españoles.

En una primera fase llevada a cabo entre febrero y mayo de 2002, se hizo un estudio de viabilidad de las tecnologías XML para la creación de un sistema informatizado de tratamiento, archivo y difusión de las resoluciones judiciales. Además, con los resultados obtenidos y su posterior difusión se perseguía impulsar dicho sistema en el ámbito de la Administración de Justicia para la edición, archivo y envío de todas las resoluciones judiciales.

Este sistema sería capaz de acometer el vaciado de información sensible de las resoluciones desde el CENDOJ, en vez de delegarlo en las empresas editoriales como se venía haciendo en el momento del contrato. También dispondría de un sistema de numeración único que permitiría identificar unívocamente todas las resoluciones archivadas en el CENDOJ, y por extensión las que fueran citadas en las mismas. Permitiría la publicación de resoluciones aisladas o por lotes y haría po-

sible seleccionar por distintos procedimientos (buscador, índice estructurado, etcétera) resoluciones concretas para que luego pudieran ser visualizadas o impresas por el usuario. También generaría lotes de resoluciones para ser suministrados, por ejemplo, a las editoriales, y permitiría obtener versiones de la resolución en diferentes formatos, por ejemplo HTML para su publicación en web, texto plano para su reutilización genérica o PDF para una impresión de calidad. Además ofrecería un servicio de búsquedas de resoluciones potente y flexible. Se podría aprovechar una estructura en campos normalizada para poder hacer búsquedas en dichos campos combinándolas en su caso con búsquedas sobre el texto estructurado de la resolución. Cabría la posibilidad de acceder a otras resoluciones (jurisprudencia) o a leyes referenciadas en una resolución dada, y obtener diferentes datos de tipo estadístico a partir de la estructura de campos normalizada. El sistema facilitaría el intercambio electrónico de documentos, así como la posibilidad de ofrecer estos servicios de una forma personalizada atendiendo a diferentes perfiles de usuario.

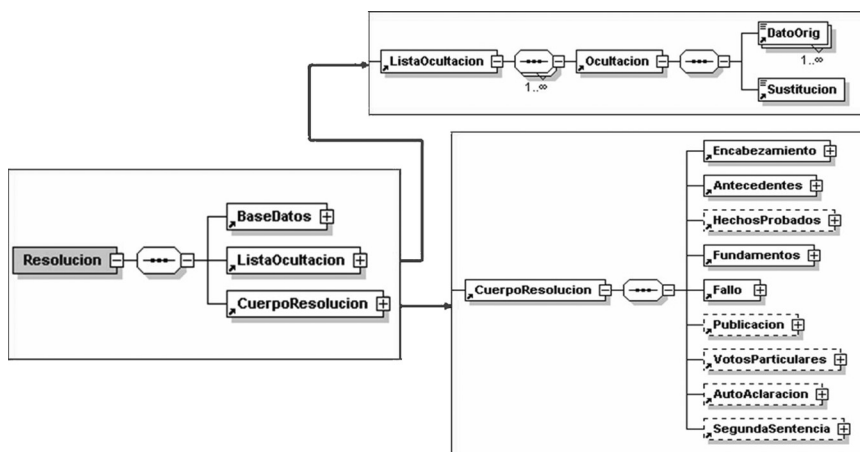


Fig. 4. Documentación del CENDOJ. Propuesta de estructura formal de una resolución.

Para ello, se proponía que el archivo electrónico de resoluciones se realizara utilizando XML, que, entre otras cosas, posibilitaría el marcado de las partes a vaciar de información sensible y permitirá realizar el vaciado automático de las resoluciones utilizando hojas de estilo. Asimismo, empleando hojas de estilo XSL sería posible realizar la conversión de la resolución a formato HTML, texto plano o PDF. El sistema único de numeración de resoluciones (tanto interno para el CENDOJ, en su primera fase, como de ámbito universal, de cara a un futuro) sería un elemento más de la aplicación XML. La estructura de campos, que viene representada en XML por una estructura de elementos, permite que el documento

XML contenga tanto el texto estructurado de la resolución como dicha estructura de campos. Por lo tanto, el documento XML contiene toda la información necesaria para poder realizar las búsquedas. La información contenida en el documento XML también se podría utilizar para la elaboración de estadísticas. El acceso a jurisprudencia y leyes relacionadas se implementaría por medio de vínculos XLink.

La segunda fase se desarrolló entre agosto y septiembre de 2002, en el marco de un nuevo contrato de asesoría de los dos departamentos antes citados de la Universidad Carlos III de Madrid y PSd, con el fin de participar en el concurso abierto por el CGPJ para la contratación del servicio de tratamiento, gestión electrónica documental, publicación y distribución de las resoluciones judiciales en el CENDOJ, publicado en el BOE de 26 de julio de 2002. A este concurso PSd acudía junto con Telefónica Sistemas y Software AG. Para ello, el equipo multidisciplinar de trabajo elaboró una DTD de XML para las resoluciones judiciales, sentencias y autos emitidos por los diversos órganos judiciales existentes en España, ajustada, en sus mínimos, a lo contemplado en el pliego de prescripciones técnicas del concurso y sus correspondientes anexos, pero contemplando como elementos opcionales otras características y circunstancias propias del tipo de documento para el que se diseñaron, con objeto de apuntar posibles mejoras en el tratamiento de los documentos previsto en el mencionado pliego. Asimismo, y formando parte también de las especificaciones del citado concurso, se procedió al marcado, ajustado a la DTD elaborada, de diez resoluciones judiciales procedentes de diversos órganos. Entre las aportaciones de esta DTD cabe destacar su capacidad para la adición de elementos no presentes en el texto de la resolución y que constituirían metainformación de gran interés, y de todos aquellos aspectos contemplados en el primer estudio de viabilidad y que no quedaron reflejados en el mencionado pliego del concurso (Modet et ál., 2002).

6. Algunas conclusiones

Las tecnologías XML resultan muy adecuadas para el tratamiento y difusión en Internet de información jurídica que, a efectos de la necesaria publicidad de las normas, supone una opción de muy bajo coste, un alcance óptimo (aunque con las limitaciones que puedan imponer a parte de la ciudadanía los costes de los ordenadores y las tarifas telefónicas, que por otro lado decrecen continuamente) y una gran sencillez para incorporar nueva información en el sistema.

Los documentos jurídicos permiten demostrar la utilidad y las ventajas de los lenguajes de marcado en general, y especialmente de XML y sus estándares acompañantes, frente a las bases de datos documentales convencionales. Entre sus principales ventajas está la capacidad de identificar en contexto los diferentes datos o porciones de información que aparecen en el texto, añadiendo marcas, sin necesidad de extraer dicha información y almacenarla fuera de contexto para que adquiera una

semántica concreta. Además permite definir estructuras documentales con el grado de flexibilidad requerido por estos documentos.

Las tecnologías XML permiten establecer cualidades que aportan semántica a los enlaces, de manera que, por ejemplo, es posible especificar el tipo de fundamento jurídico en el que se basa la relación entre varios textos jurídicos.

La labor de recopilación de toda la normativa que regula una materia concreta también se facilita gracias a las posibilidades de reutilización de información que ofrecen las tecnologías XML para la generación de documentos históricos (todas las redacciones) y vigentes (las redacciones vigentes en un momento dado) a partir de la reutilización de diferentes partes de distintas normas. Así mismo se facilita la elaboración de índices de acceso por los datos identificativos tanto para la normativa como para las resoluciones judiciales.

El uso de XML permite establecer múltiples formas de presentación de un mismo documento, tanto en lo referente al formato electrónico concreto (HTML, PDF, etcétera) como en lo tocante al contenido del mismo (ocultación de datos, versiones reducidas de sentencias...).

La colaboración entre juristas, informáticos y documentalistas es muy conveniente para conseguir sistemas que sean capaces de implementar todas las posibilidades que conlleva el tratamiento de los textos jurídicos con tecnologías de marcado, en especial XML.

Es necesaria la normalización de tipos documentales al nivel más alto posible, de forma que la estructura, tanto de normas como de resoluciones, sea definida por una misma DTD en cualquier sistema jurídico documental, como sucede, por ejemplo, con la legislación italiana.

Referencias

- Aguiló, Josep (1988). *Informática jurídica, lenguajes documentales y técnica legislativa*. Tesis doctoral. Alicante: Universidad, 1988.
- Arellano, M.^a del Carmen (2003). *La organización hipertextual del ordenamiento jurídico por medio de tecnologías XML*. Tesis Doctoral. Madrid: Universidad Carlos III, 2003.
- Arellano, M.^a del Carmen (2005). *Aportaciones de la técnica legislativa y XML a la informática jurídica documental*. Madrid: Boletín Oficial del Estado, 2005.
- Arellano, M.^a del Carmen.; Caridad, Mercedes; Hernández, Antonio; Martín, Bonifacio; Moreira, J. Antonio; Nogales, J. Tomás; Robledano, Jesús; Rodríguez, David (2001). *Desarrollo de sistemas de información hipertextuales para la difusión y el acceso a la legislación*. Póster. // IFLA Council and General Conference Libraries and Librarians: Making a Difference in the Knowledge (67.th, Boston, 2001). Panel n.º 51.
- Arellano, M.^a del Carmen; Nogales, J. Tomás; Martín, Bonifacio (2003). *La organización hipertextual del ordenamiento jurídico: posibilidades de XML y estándares relacionados*. // *Revista general de información y documentación*. 13:2 (2003) 181-191.

- Cazorla Prieto, Luis María (1999). *Codificación contemporánea y técnica legislativa*. Pamplona: Aranzadi, 1999.
- Martín, Bonifacio (2002). *Tratamiento y difusión en Internet de información jurisprudencial mediante tecnologías XML: aplicación al caso de Tribunal Constitucional*. Tesis Doctoral. Madrid: Universidad Carlos III, 2002.
- Modet, José M.ª; Corrales, Pablo; Sánchez, Luis; Nogales, J. Tomás; Martín, Bonifacio; Arellano, M.ª del Carmen (2002). *Estudio de viabilidad para la implantación y normalización de un sistema basado en el estándar XML para la gestión de resoluciones*. Informe técnico conjunto entre PSD Gestión Documental, S. L., y la Universidad Carlos III de Madrid (Dptos. Biblioteconomía y Documentación / Ingeniería Telemática) para el Centro de Documentación Judicial (CENDOJ) del Consejo General del Poder Judicial (CGPJ) (mayo-julio de 2002).
- Hernández Marín, R. (1998). *Introducción a la teoría de la norma jurídica*. Madrid: Marcial Pons, 1998.
- López-Muñiz Goñi, Miguel (1984). *Informática jurídica documental*. Madrid: Díaz de Santos, 1984.
- Martín Casals, Miquel (1989). *Preámbulo y disposiciones directivas*. // GRETEL (Grupo de Estudios de Técnica Legislativa). *Curso de técnica legislativa*. Madrid: Centro de Estudios Constitucionales, 1989. 73-101.
- Nogales, J. Tomás; Arellano, M.ª del Carmen (2000a). *La organización hipertextual de textos legislativos con HTML y XML: una necesidad y las soluciones de presente y futuro*. // *Actas de las VII Jornadas Españolas de Documentación FESABID 2000*. Bilbao: FESABID, 2000. 385-393.
- Nogales, J. Tomás; Martín, Bonifacio; Arellano, M.ª del Carmen; Rodríguez, David; Hernández, Antonio; Caridad, Mercedes; Moreiro, J. Antonio; Robledano, Jesús (2000b). *La difusión de los textos legislativos en Internet haciendo uso de los lenguajes de marcado HTML y XML*. // *Actas de la III Conferencia Internacional de Derecho e Informática*. La Habana: Ministerio de la Informática y las comunicaciones, 2000. CD-ROM.
- Nogales, J. Tomás; Arellano, M.ª del Carmen; Rodríguez, David; Martín, Bonifacio; Hernández, Tony (2000c). *Un repertorio legislativo hipertextual mediante marcado de texto: las disposiciones generales del Boletín Oficial de la Comunidad de Madrid*. // *JBIDI 2000*. Primeras Jornadas de Bibliotecas Digitales (Valladolid, 6 y 7 de noviembre). Valladolid: Universidad, 2000. 89-104.
- Nogales, J. Tomás; Arellano, M.ª del Carmen; Martín, Bonifacio (2002). *Modelos formales para la definición estructural y semántica en documentos XML. Comparación de posibilidades en un corpus textual de documentación jurisprudencial*. // *García, Purificación; Canós, José Hilario (coords.). III Jornadas de Bibliotecas Digitales (JBIDI'2002)*. Madrid: Universidad Politécnica, 2002. 97-106.
- Nogales, J. Tomás; Martín, Bonifacio; Arellano, M.ª del Carmen (2003a). *Informática, derecho y documentación. Experiencias y posibilidades de aplicación de los lenguajes de marcado de texto (SGML, HTML y XML) a los documentos jurídicos*. // *Dávara, Miguel (coord.)*. *Encuentro sobre Informática y Derecho*. (16.º, Madrid, 2002). Madrid: Instituto de Informática Jurídica, Universidad Pontificia de Comillas, 2003. 335-374.

- Nogales, J. Tomás; Martín, Bonifacio; Arellano, M.^a del Carmen (2003b). Una propuesta para el tratamiento documental de las resoluciones judiciales en España haciendo uso de tecnologías XML. // Los sistemas de información en las organizaciones: eficacia y transparencia. FESABID 2003: 8.^{as} Jornadas Españolas de Documentación (Barcelona, 6, 7 y 8 de febrero de 2003). Barcelona: COBDC, 2003. 385-393.
- Nogales, J. Tomás; Arellano, M.^a del Carmen; Martín, Bonifacio (2004). Propuesta de aplicación de los criterios de técnica legislativa a un sistema de información de legislación usando tecnologías XML. // XVI Encuentro sobre Informática y Derecho. Madrid: Instituto de Informática Jurídica, 2004. 273-287.
- Otto, Ignacio de (1988). Derecho constitucional: sistema de fuentes. Barcelona: Ariel, 1988.
- Rodríguez-Zapata, Jorge (1987). Sanción, promulgación y publicación de las leyes. Madrid: Tecnos, 1987.
- Sainz Moreno, Fernando (1993). La publicidad de las normas. // Martín-Retortillo Baquer, Lorenzo (coord.). La protección jurídica del ciudadano (procedimiento administrativo y garantía jurisprudencial): estudios en homenaje al profesor Jesús González Pérez. Madrid: Civitas, 1993. 121-146, esp. 125.
- Sánchez Morón, Miguel (1989). Contenido de las normas, principio de homogeneidad, estructura formal. // Sáinz Moreno, Fernando; Da Silva Ochoa, Juan Carlos (coords.). La calidad de las leyes. Vitoria: Parlamento Vasco, 1989. 101-119.
- Viver Pi-Sunyer, Carles (1989). La parte final de las leyes. // GRETEL (Grupo de Estudios de Técnica Legislativa). Curso de técnica legislativa. Madrid: Centro de Estudios Constitucionales, 1989. 135-172.

Recibido: 2008-05-02. Aceptado: 2008-09-02