

Tratamiento de resoluciones judiciales

Cliente

Nombre

Centro de Documentación Judicial (CENDOJ) - Consejo General del Poder Judicial.

Sector

Administración Pública - Justicia.

Perfil del cliente

El CENDOJ tiene entre sus funciones la recopilación y difusión de la jurisprudencia de los tribunales en virtud de la competencia que a tal fin otorga al Consejo General del Poder Judicial el artículo 107.1 de la Ley Orgánica del Poder Judicial.

Fase 1

Datos básicos del Proyecto

Situación

Hasta 1999, el CENDOJ disponía de un procedimiento para la entrega de las resoluciones judiciales a las editoriales jurídicas consistente en la distribución de CR-ROM conteniendo las imágenes procedentes de la digitalización de los papeles originales de las resoluciones.

Objetivos

El objetivo del proyecto consiste en mejorar la base de datos documental de jurisprudencia obteniendo el texto de las resoluciones, incrementando el número de datos asociados a las mismas y mejorando los sistemas de consulta y distribución hacia los clientes y usuarios del CENDOJ.

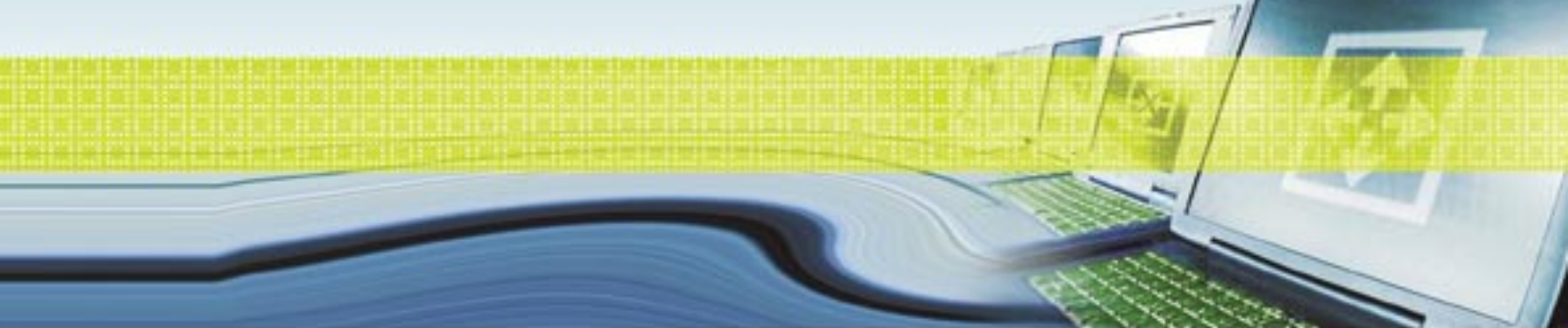
Solución

La solución adoptada por SERIKAT para conseguir los objetivos del proyecto se ha basado en dos elementos fundamentales. Por una parte, se ha definido un procedimiento de trabajo para el tratamiento de las resoluciones que atiende tanto a la productividad del proceso como a la calidad del resultado. Por otra parte, en este proceso se han integrado varias herramientas comerciales que permiten digitalizar papel, obtener su texto mediante técnicas de reconocimiento óptico, extraer datos de los documentos y generar las bases de datos documentales que se entregan al cliente.

Beneficios

- ➔ Se agiliza la distribución de las resoluciones a las editoriales jurídicas.
- ➔ Las editoriales jurídicas pueden realizar búsquedas flexibles para localizar las resoluciones de su interés.
- ➔ Otros usuarios en el CENDOJ pueden realizar la explotación de la base de datos global de jurisprudencia.





Resumen del Proyecto

SERIKAT, como adjudicatario del concurso convocado por el Consejo General del Poder Judicial (CGPJ) en julio del año 1999, ha realizado un proceso de tratamiento documental de resoluciones judiciales (autos y sentencias) desde dicha fecha hasta el mes de octubre del año 2002.

El objetivo perseguido por el CGPJ con este proyecto se centraba en la creación de una base de datos documental con las resoluciones judiciales y los datos a ellas asociadas para que el Centro de Documentación Judicial (CENDOJ) dispusiera de una herramienta informática ágil y flexible que permitiera el suministro de las resoluciones a sus usuarios tanto internos como externos.

Las resoluciones originales se encontraban en tres formatos: papel, ficheros ofimáticos de tipo texto (Word, RTF, TXT...) y ficheros de imagen TIFF. Para el tratamiento de estas resoluciones se puso en marcha un flujo de trabajo que era aplicado a la totalidad de las resoluciones aunque con tratamientos distintos en función del formato de entrega. Las operaciones principales que componían este flujo de trabajo eran las siguientes:

- ➔ Digitalización del papel.
- ➔ Tratamiento de la imagen para aumentar su calidad y mejorar el nivel de reconocimiento óptico.
- ➔ Reconocimiento óptico de caracteres (OCR) de la imagen resultante.
- ➔ Corrección del texto obtenido del reconocimiento hasta obtener un texto totalmente fiel al original.
- ➔ Conversión de las resoluciones con una estructura homogénea.
- ➔ Obtención del conjunto de datos principales de cada resolución.
- ➔ Creación de una base de datos documental, soportada por el gestor ALCHEMY, en la que se unen las imágenes, el documento Word normalizados y los datos de cada resolución.
- ➔ Entrega de las resoluciones tratadas.

Durante el período de prestación del servicio se procesaron la totalidad de las resoluciones generadas en dicho período por todas las Audiencias Provinciales, todos los Tribunales Superiores de Justicia de las Comunidades Autónomas, la Audiencia Nacional y el Tribunal Supremo, así como un conjunto de ellas generadas anteriormente por estos mismos tribunales.



Fase 2

Datos básicos del Proyecto

Situación

Hasta 2002, el CENDOJ disponía de una base de datos documental con el texto completo de las resoluciones judiciales y con un conjunto de datos asociados a las mismas.

Objetivos

El objetivo del proyecto consiste en mejorar la base de datos documental de jurisprudencia incrementando el número de datos asociados a las mismas, localizando diversos elementos en el propio texto de las resoluciones (secciones, referencias de legislación y jurisprudencia...) y realizando la ocultación de información sensible de carácter personal.

Solución

La solución diseñada por SERIKAT para este proyecto ha girado en torno a dos elementos principales. El primero de ellos ha sido la definición de un modelo de resolución judicial en formato XML que permite recoger todos los elementos del tratamiento. El segundo elemento ha sido el diseño de un servicio de tratamiento de resoluciones gobernado por un sistema informático diseñado y desarrollado por SERIKAT a la medida de las necesidades concretas del proyecto.

Resumen del Proyecto

En el mes de octubre del año 2002, SERIKAT resultó adjudicataria del concurso convocado por el CGPJ para realizar nuevos tratamientos sobre las resoluciones judiciales y para implantar una solución de distribución electrónica de las mismas, que tendría efecto en la prestación del servicio a partir del mes de octubre de 2002.

El objetivo de este nuevo proyecto puede ser considerado como una perfección del anterior, ejecutado desde julio de 1999. Las mejoras introducidas en el objetivo persiguen, en primer lugar, ofrecer a las editoriales jurídicas las resoluciones judiciales libres de información sensible de carácter personal, cumpliendo así con el requisito legal de protección de información. Por otra parte, también se considera objetivo de este proyecto la introducción de tecnologías que faciliten y automaticen la relación de intercambio comercial de las resoluciones entre el CENDOJ y las editoriales jurídicas.

Beneficios

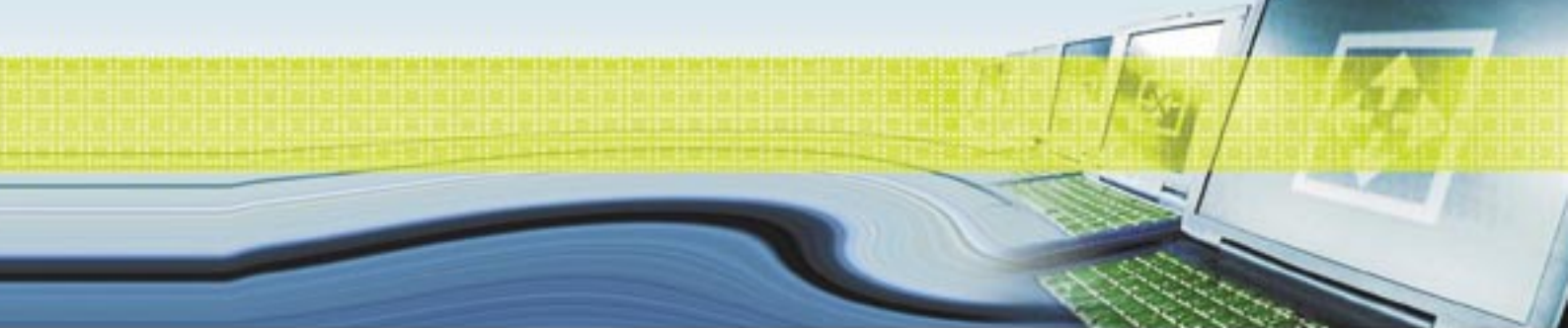
- ➔ El formato XML mejora la explotación documental de los datos y documentos.
- ➔ Los documentos pueden ser presentados con o sin información sensible de carácter personal en función del perfil del usuario que los consulta.
- ➔ La localización de elementos en el texto permite su utilización en la explotación del documento.



Las resoluciones son, inicialmente, suministradas en tres formatos: papel, ficheros ofimáticos de tipo texto (Word, RTF, TXT...) y bases de datos LOTUS.

El tratamiento que se aplica a las resoluciones en este servicio, además de operaciones similares a las aplicadas en el servicio anterior, incluye nuevas operaciones de una relevancia muy significativa entre las que cabe señalar las siguientes:

- ➔ **Ocultación de información sensible de carácter personal.** Todos los datos (nombres, direcciones, números de identificación...) presentes en las resoluciones que puedan servir para la identificación de personas físicas que sean citadas como partes (demandante, demandado, apelante, apelado...), testigos, peritos, etcétera, son sustituidos por datos alternativos que oculten su personalidad. El resultado de este tratamiento es un documento en



formato XML que, aplicando distintas plantillas (XSL) puede ser presentado con los datos originales o con la información ocultada en función de los privilegios o derechos de consulta de los usuarios que acceden a las resoluciones.

⇒ **Identificación de la estructura de la resolución.** La estructura de las resoluciones es analizada para identificar sus partes estándar: cabecera, antecedentes de hecho, fundamentos de derecho, parte dispositiva, votos particulares y segundas sentencias. Estas partes son “marcadas” en la estructura XML que se genera, permitiendo que, en función de

los privilegios o derechos de consulta de los usuarios, pueda mostrarse la totalidad de la resolución o exclusivamente alguna de sus partes.

⇒ **Obtención final de las resoluciones en formato XML.** El resultado final del tratamiento queda recogido en un archivo XML que se divide en tres secciones: la información que va a formar parte de la base de datos; las cadenas de ocultación que enlazan la información original con la ocultada y su posición en el texto; y el cuerpo de la resolución identificando cada sección y cada párrafo.

Fase 3

Como se ha comentado en la sección “En marcha”, SERIKAT ha sido recientemente adjudicataria del nuevo concurso para el tratamiento de las resoluciones judiciales convocado por el Consejo General del Poder Judicial. La clave perseguida por el CENDOJ en esta nueva fase de trabajo puede ser resumida en una sola palabra: flexibilidad.

⇒ Flexibilidad para permitir la introducción en el sistema de nuevas fuentes (las Administraciones de Justicia de Comunidades Autónomas) de resoluciones con formatos distintos.

⇒ Flexibilidad para cambiar los procesos productivos ante el cambio de especificaciones de las resoluciones originales.

⇒ Flexibilidad para que cada resolución siga caminos productivos distintos en función del estado en el que entra al sistema y, por lo tanto, en función de las tareas a las que tiene que ser sometida.

⇒ Flexibilidad en la capacidad de trabajo para asumir la variación del número de resoluciones que se introducen en el sistema productivo por la incorporación de nuevos tribunales como origen de documentos.

Este nuevo escenario ha introducido una mayor complejidad en el mapa de procesos productivos en los que se divide el proyecto. Por ello, SERIKAT ha realizado un esfuerzo en la remodelación de su sistema productivo para dotarlo de la capacidad de asumir estos nuevos escenarios y, adicionalmente, para dotarlo de la flexibilidad de cambio a lo largo del tiempo.

*Jose Ignacio Goicoechea,
Director de Proyectos*

