

# **Utilización de la web semántica y las ontologías para la difusión, el trabajo y la investigación sobre textos normativos: la experiencia del Estatuto de Autonomía de Aragón**

Jesus Divasson Mendivil (Gobierno de Aragón)

José Luis Bermejo Latre, Javier Garcia Marco (Laboratorio Jurídico-Empresarial, Univ. de Zaragoza)

Antonio C. Calvo de la Prida y Pedro Brusel Andrés (Iritec)

## **Resumen**

El Gobierno de Aragón ha desarrollado una página dedicada al Estatuto de Aragón en la que recursos de distintas clases relacionados con el mismo se organizan utilizando técnicas de web semántica. El web incluye una intranet con la cual un equipo de colaboradores formado por expertos trabaja de forma colaborativa para elaborar los contenidos del web, y diversas herramientas participativas, dirigidas a la difusión del Estatuto. Pero, la zona más innovadora del web es la destinada a la lectura del Estatuto ya que, mediante la elaboración de una ontología específica y el etiquetado XML del texto, se facilita al usuario el acceso contextual a los recursos relacionados con cada parte del mismo.

## **1 Introducción**

La Vicepresidencia del Gobierno de Aragón tiene atribuida dentro de sus competencias generales el impulso, análisis y coordinación de la política de desarrollo estatutario. Bajo la dependencia de la Vicepresidencia se encuentra la Dirección General de Desarrollo Estatutario a la que le corresponde en el ejercicio de sus funciones fomentar la investigación sobre la evolución del Estado autonómico y divulgar el conocimiento del Estatuto de Autonomía de Aragón y su desarrollo mediante la edición de publicaciones, la organización de jornadas, seminarios y otras formas de difusión.

El nuevo Estatuto de Autonomía de Aragón supone un impulso en el desarrollo de la autonomía aragonesa y se hace preciso dar a conocer a la sociedad aragonesa las características de la nueva norma y crear los mecanismos necesarios para permitir a los juristas profundizar en la interpretación de las reformas efectuadas.

En la consecución de esta finalidad la Dirección General de Desarrollo Estatutario creó la página web "[www.estatutodearagon.es](http://www.estatutodearagon.es)", con la pretensión de ofrecer un marco divulgativo y formativo sobre Aragón y su Estatuto de Autonomía. Pero se trataba de una página estática en la que únicamente es posible consultar el texto de algunas normas y acceder a informes y presentaciones sobre el Estatuto.

Se tomó por ello la decisión de desarrollar un nuevo sitio web dedicado al Estatuto de Autonomía de Aragón, que permita aprovechar al máximo las funcionalidades de la web para hacer de Internet un medio para la difusión del Estatuto de Autonomía de Aragón, y también un instrumento de trabajo y comunicación para los expertos interesados en su estudio y análisis.

Para el desarrollo de la nueva página web se formó un Comité, integrado por Xavier de Pedro Bonet, Director General de Desarrollo Estatutario, Jesús Divasson Mendivil y Francisco Doblás Gálvez. La realización del análisis funcional y la elaboración de la ontología se encargó al Laboratorio Avanzado de la Universidad de Zaragoza sobre Aplicaciones Jurídicas y Empresariales en la Sociedad de la Información, situado en el Parque Tecnológico Walqa (Huesca), cuyo equipo estuvo formado por José Félix Muñoz Soro, Francisco Javier García Marco y José Luis Bermejo Latre. Del desarrollo de la página web se encargó la empresa Iritec, cuyo equipo estuvo constituido por Iván Toda Caraballo, Jefe de Proyecto, Pedro Brusel

Andrés y Antonio Calvo de la Prida. Finalmente, una vez desarrollada la página un equipo de juristas, dirigido por el catedrático de la Universidad de Zaragoza, José Bermejo Vera, se encarga de elaborar sus contenidos.

El elemento más característico del proyecto ha sido la utilización de las técnicas de la web semántica, con una doble finalidad:

- Proporcionar una representación del conocimiento del Estatuto de Aragón en formatos estándar, con el objetivo de que otras webs y aplicaciones del Gobierno de Aragón o de terceros puedan aprovechar computacionalmente el conocimiento generado dentro del web.
- Lograr una estructuración de contenidos que permita implementar un conjunto de algoritmos internos en el web destinados a lograr una organización de contenidos más relevante. Esto se plasma en las estructuras de navegación y en los procesos de búsqueda. De esta manera, mediante deducciones lógicas sobre el conocimiento generado, el web "entiende" el significado del texto del estatuto y de todos los contenidos expuestos y puede guiar al usuario y recomendarle nuevos contenidos de interés.

## **2 El sitio web del Estatuto de Aragón**

### **2.1 Planteamiento**

El sitio web dedicado al Estatuto de Aragón se divide en varias zonas, correspondiendo cada una ellas a un objetivo concreto:

- **Divulgación:** Esta zona contiene materiales divulgativos relacionados con el Estatuto. Es la que se corresponde con los objetivos de la página existente previamente, aunque se mejoran notablemente tanto el diseño como los contenidos.
- **Noticias:** En este apartado se publican todas las noticias que se generan relacionadas con el Estatuto y con su desarrollo, bien sean acciones del Gobierno de Aragón o de otras instituciones. La pretensión es acompañar cada una de estas noticias con recursos audiovisuales como videos, audios y fotografías. Esta sección se actualiza diariamente por personal del Gabinete de Comunicación del Gobierno de Aragón.
- **Investigación:** Tiene dos partes, una es la intranet destinada al equipo de colaboradores, que son juristas y otros expertos que elaboran materiales sobre el Estatuto y que disponen en la misma de herramientas colaborativas. Una vez elaborados por el equipo de colaboradores, los materiales sobre el Estatuto se publican en la zona de investigación del web, que es de acceso público.
- **Lectura del Estatuto:** en esta zona se puede leer el Estatuto y se dispone de accesos contextuales a definiciones de conceptos, documentos y otros recursos del web.

Una característica del web es que todas las zonas comparten un único repositorio de documentos en el que, además de los métodos de indexación tradicionales y las búsquedas por texto completo, se utilizan las técnicas propias de la web semántica para la realización de búsquedas.

También se dispone en el web de herramientas participativas que pueden ser incorporadas en cualquiera de las zonas y que son: cuestionarios, votaciones, foros, chat (su finalidad es establecer diálogos entre los usuarios y una determinada persona, como un colaborador o un responsables político), editor wiki y calendario. Cada una de las instancias de estas herramientas, como por ejemplo una votación, un foro temático o un documento que esté siendo editado en el editor wiki, puede

ser etiquetada y buscada de forma equivalente a los documentos, utilizando las categorías de la ontología. Ello permite que el usuario acceda a las mismas a través de las opciones de búsqueda y también que puedan ser incorporadas en los accesos contextuales de la zona de lectura.

## **2.2 Zona de investigación**

El objetivo principal de esta zona es ser una herramienta de trabajo para los investigadores y estudiosos interesados en el Estatuto de Autonomía de Aragón. El medio principal para ello es la publicación en el web de una amplia selección de documentos sobre el Estatuto, recopilados por un grupo de expertos, a los que denominamos "colaboradores"<sup>1</sup>.

Además, mediante la inclusión de herramientas participativas, se persigue la creación de una comunidad en la que todas las personas interesadas en el Estatuto de Autonomía de Aragón puedan trabajar y dialogar sobre el mismo desde diferentes lugares, posiciones, intereses y tiempos, aprovechando las opciones que proporciona la página web.

Para elaborar los recursos, los colaboradores disponen de una intranet en la que se les proporcionan los medios para generar nuevos materiales mediante trabajo colaborativo y en la que también se facilita la comunicación y el debate entre ellos. Así, la intranet permite que los colaboradores intercambien información durante la fase de elaboración y/o preparación de los contenidos originales, aporten determinados contenidos para la previa valoración por el grupo de su interés e incluso que redacten de modo colaborativo algunos textos, utilizando un wikieditor. Otra función muy importante de la intranet es el mantenimiento de la ontología por parte de los colaboradores.

Finalmente también pueden incorporarse en la intranet herramientas participativas, que faciliten la creación de comunidad a la pequeña escala del equipo de colaboradores.

## **2.3 Zona de lectura del Estatuto**

La finalidad de esta zona es que el usuario pueda leer el texto completo del Estatuto, pero dentro de un entorno que le permite el acceso inmediato a los recursos directamente vinculados a la parte de la norma que esté leyendo en ese momento. Aquí es donde se aprovechan al máximo todas las posibilidades que proporcionan la ontología y la indexación en base a la misma de todos los recursos incorporados al web.

Como base de esta operatoria, el Estatuto de Aragón se etiqueta utilizando un dialecto del XML desarrollado al efecto. Para ello, cuando una palabra o expresión se corresponde con una de las categorías de la ontología se marca con la correspondiente etiqueta, que la relacionará con la misma. Así se permite que las búsquedas y el acceso a los recursos puedan realizarse directamente a través de enlaces contenidos en el propio texto, con el interface que se describe a continuación.

En la pantalla de lectura las palabras o expresiones etiquetadas aparecen destacadas en negrita y cambian de color cuando se pasa el cursor por encima de

---

<sup>1</sup> El equipo de colaboradores, coordinado por el profesor José Bermejo Vera, de la Universidad de Zaragoza, está formado por un grupo de especialistas en diversas parcelas del Derecho público (constitucional, administrativo y fiscal) y privado (civil y foral), pudiendo incorporarse también a especialistas en economía y en historia. El equipo se encarga de redactar contenidos originales, recopilar y seleccionar documentos y materiales de interés relativos al Estatuto de Autonomía.

las mismas. Si el usuario hace click sobre ellas inmediatamente se realiza una búsqueda utilizando la categoría relacionada. Una vez finalizada la búsqueda únicamente se presentan al usuario, en la columna con el texto del Estatuto, los artículos del mismo que están relacionados con la categoría seleccionada. Las palabras o expresiones coincidentes con la utilizada en la búsqueda se destacan mediante texto inverso. Un botón ubicado en la parte superior permite en todo momento desactivar la selección del texto y volver a leer el Estatuto completo.

También se presenta al usuario otra información relacionada con la categoría o categorías de búsqueda en las "cajas" situadas a la derecha del texto del Estatuto. En la primera de ellas se muestran los recursos asociados, que pueden ser de cualquiera de los tipos contemplados en el web. El usuario puede acceder al contenido de los recursos clicando sobre su título. En la segunda "caja" se indican los conceptos de la ontología asociados a la categoría utilizada en la búsqueda, mostrando el superior en la jerarquía y, si los hay, los que comparten el mismo nivel bajo este y los que se encuentran bajo la categoría seleccionada. Si el usuario hace clic sobre estas categorías se realiza una nueva búsqueda.

A la izquierda de la zona de lectura se sitúa otra columna, que puede desplegarse u ocultarse a elección del usuario. Esta columna permite, en primer lugar, navegar por el texto del Estatuto utilizando el índice sistemático. La segunda opción que contiene la columna es un índice alfabético, que incluye todas las categorías de la ontología ordenadas alfabéticamente. Si se selecciona una de ellas se realiza la búsqueda de la misma forma que cuando se clica sobre la categoría en el texto del Estatuto.

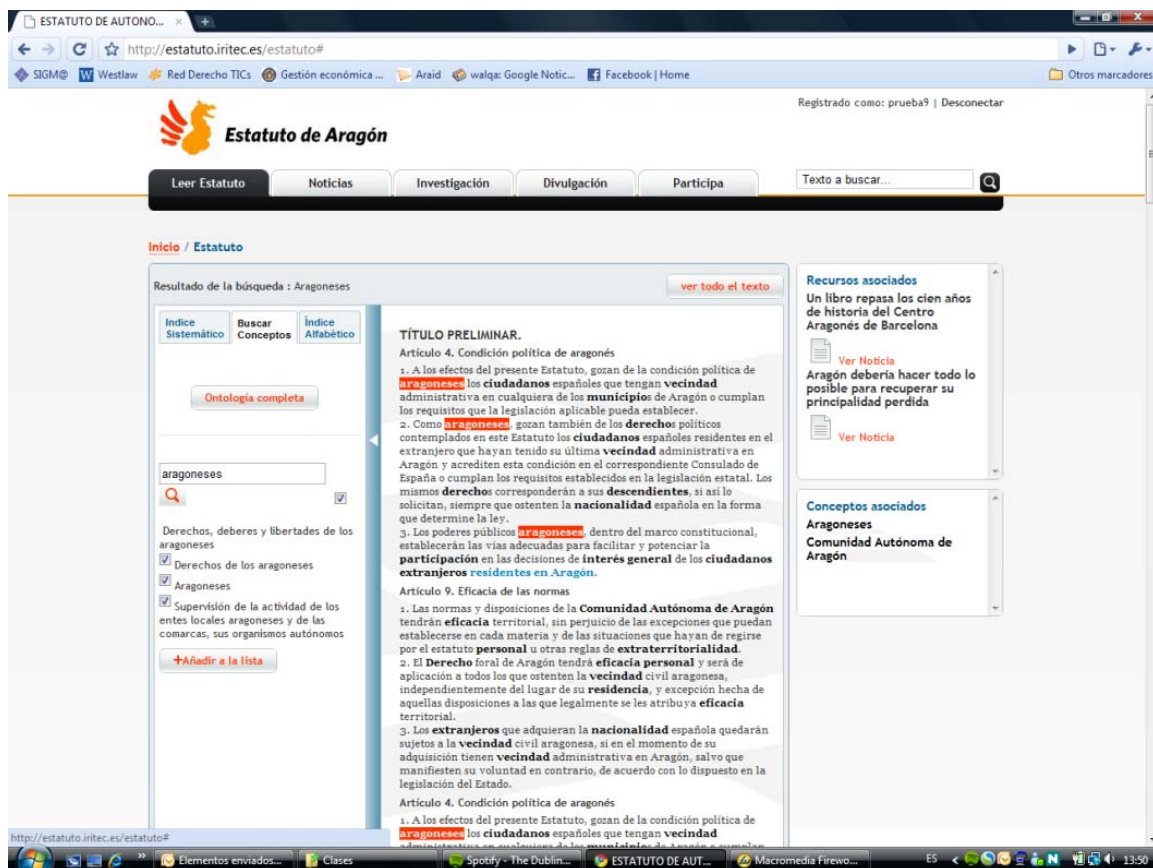


Fig. 1. Zona de lectura del Estatuto.

La tercera opción es la búsqueda sobre la ontología, que puede realizarse de dos maneras. La primera es acceder al árbol completo de la ontología, por el cual se puede navegar desplegando y plegando las categorías hasta llegar a la que se quiere. En la segunda modalidad, el usuario introduce un texto y el sistema busca todas las categorías relacionadas con el mismo, presentando la lista resultante al usuario. Este puede seleccionar las categorías que desea mediante un check box situado junto a cada una de ellas y luego las añade a la lista de búsqueda, mediante la correspondiente opción.

La operación anterior puede repetirse tantas veces como el usuario desee y permite componer una lista de búsqueda que integre diferentes partes de la ontología. Al lanzarse la búsqueda el sistema seleccionará y presentará todos los artículos relacionados con las categorías presentes en la lista de búsqueda y otro tanto hará con los recursos.

## **2.4 Los recursos**

Los recursos del web incluyen, por una parte, contenidos divulgativos y, por otra, contenidos especializados, elaborados por el equipo de colaboradores y destinados a la comunidad especializada en la interpretación y desarrollo del Estatuto, que incluye a estudiantes, investigadores, políticos y profesionales que precisen un conocimiento más o menos profundo de los diferentes aspectos relativos al desarrollo del Estatuto de Autonomía.

Los contenidos especializados del web pueden ser de los siguientes tipos:

- **Voces:** Artículos explicativos de los diferentes conceptos que componen la ontología. Con una extensión de hasta 5.000 caracteres, se muestran en formato HTML.
- **Conceptos:** Descripciones de los conceptos jurídicos relacionados con el Estatuto. Su extensión es de hasta 20.000 caracteres y se muestran también en formato HTML.
- **Legislación:** Incluye legislación supranacional, estatal y autonómica.
- **Jurisprudencia:** Incluye resoluciones judiciales que interpretan y aplican normas de desarrollo del Estatuto aragonés (u otros, aunque principalmente de éste). En particular, sentencias del Tribunal Constitucional, del Tribunal Supremo, del Tribunal Superior de Justicia de Aragón y de los de otras Comunidades Autónomas, y de Juzgados de los órdenes civil y contencioso-administrativo.
- **Doctrina:** Incluye estudios académicos y doctrinales publicados en revistas de las áreas de Derecho constitucional, administrativo, civil y financiero, y economía aplicada, así como reseñas de la producción bibliográfica relacionada con el desarrollo y aplicación del Estatuto aragonés.
- **Informes y dictámenes:** Incluye informes, dictámenes o estudios encargados por cualquier ente u órgano dependiente del Gobierno de Aragón, documentos preparatorios de normas, informes o dictámenes de instituciones consultivas y documentos generados por otras instituciones autonómicas de relevancia.

Los recursos de los cuatro últimos tipos provienen de fuentes externas de acceso universal y general conocimiento, pero el valor añadido lo ofrece su filtrado, recopilación, sistematización y vinculación a los conceptos de la ontología. Gracias a esta segunda clase de contenidos, la herramienta se convierte además en un ejemplo de reutilización de la información del sector público, tal y como dispone la Ley 37/2007, de 16 noviembre.

Los recursos pueden publicarse en el web agrupados según su tipo. Pero, además, independientemente del tipo al que pertenezcan, a los recursos se les puede asignar una categoría, que se utiliza también para su publicación en el web, ya que

los contenidos pueden organizarse en "cajas", que se corresponden con las distintas categorías. Ello permite publicar los recursos agrupados según temas que los colaboradores pueden administrar como deseen.

### **3 Utilización de la web semántica**

#### **3.1 Planteamiento**

Desde el principio, se planteó como un objetivo clave del proyecto que el sistema de información sobre el Estatuto de Autonomía de Aragón tuviera un sistema de recuperación y navegación sólido y riguroso. Esto implicaba que debía solventar los problemas bien conocidos de la búsqueda de palabras clave como los silencios en los casos de sinonimia y de relaciones jerárquicas (conceptos genéricos y específicos), las ambigüedades en el caso de la polisemia y la falta de relaciones para recuperar conceptos relacionados, pero de mucha importancia para la consulta.

Este aspecto resulta clave en sistemas pequeños en los que no se puede delegar en la redundancia de documentos e informaciones para soslayar los problemas planteados. También es fundamental en sistemas especializados dirigidos a expertos, donde cada silencio es una fuente de error potencial, y donde cada falso positivo supone perder un tiempo precioso.

Además, el sistema tenía también que permitir navegar de forma jerárquica y asociativa el Estatuto y modular las búsquedas extractivas de palabras clave. Debía funcionar de forma transparente al usuario, de manera que éste no tuviera que interaccionar con las herramientas de búsqueda avanzada si no lo deseaba, pero que tuviera retroalimentación y acceso a las mismas si lo consideraba necesario o conveniente.

#### **3.2 La web semántica**

Para poder conseguir todo lo anterior, el proyecto de desarrollo de la nueva web del Estatuto de Aragón se construye sobre un núcleo tecnológico basado en los estándares y recomendaciones que componen la denominada Web Semántica. La Web Semántica es una iniciativa de la organización World Wide Web Consortium (W3C) destinada a lograr modelar una red de datos, organizada y normalizada, en contraposición a la actual web, compuesta por millones de núcleos dispersos de información.

El objetivo es, por lo tanto, establecer la normativa que permita representar de una manera formal la información disponible en las webs. Para ello W3C ha ido elaborando un conjunto de lenguajes reglados cuya finalidad es complementar al conocido HTML para categorizar y aportar "significado" a la información publicada en la Web. Al estandarizar los lenguajes dicho "significado" puede ser "entendido" e "interpretado" por otros programas y webs, permitiendo un nuevo grado de procesamiento de la información por parte de las aplicaciones computacionales, logrando así nuevas funcionalidades de interés para los usuarios.

Los estándares o recomendaciones más importantes descritos por W3C para la web semántica son:

- Resource Description Framework (RDF), el cual especifica la sintaxis para describir objetos o entidades de información. RDF permite asociar a dichos objetos propiedades y sus correspondientes valores, permitiendo de esta manera definir datos estructurados. Como formato de representación se emplea el lenguaje XML, en la sintaxis denominada RDF/XML.

- Resource Description Framework Schema (RDFS) es una extensión de RDF que incorpora la posibilidad de describir relaciones básicas entre los objetos, para permitir una representación de relaciones entre los mismos, o lo que es lo mismo, una representación del conocimiento.
- Ontology Web Language (OWL), es un lenguaje construido sobre RDFS para modelar representaciones del conocimiento más sofisticadas. Mediante OWL se pueden describir clases e individuos que las componen. Para cada clase se permiten especificar un conjunto de aseveraciones y relaciones sofisticadas entre cada uno de las entidades. OWL permite el modelado conceptual de cualquier campo de conocimiento humano. A partir de esta información se pueden realizar inferencias y deducciones sobre las entidades descritas, dotando así de "inteligencia" a las aplicaciones computacionales.

### **3.3 Estándares utilizados**

La solución adoptada en el proyecto residió, lógicamente, en la utilización de un vocabulario estructurado para la recuperación de la información ISO/DIS 25964 que permitiera resolver sinónimos y términos polisémicos, ampliar y reducir búsquedas de forma automática basándose en la proximidad conceptual, presentar jerárquicamente la estructura conceptual del Estatuto para facilitar su navegación, así como recuperar de forma deliberada conceptos relacionados estrechamente con el que se estuviera buscando y relevantes para la necesidad de información expresada.

El vocabulario estructurado se ha diseñado como un tesoro facetado y se ha implementado como una ontología en OWL mediante el estándar SKOS, de manera que fuera utilizable en un entorno de web semántica. SKOS es un caso práctico de aplicación de OWL destinado a modelar fácilmente conjuntos de conceptos y sus relaciones de parentesco, sinonimia o asociación existentes. Siguiendo estas reglas se puede modelar un árbol conceptual extenso de todos los términos relacionados con el Estatuto de Autonomía de Aragón.

La utilización de estándares de web semántica en la ontología —SKOS, OWL y RDFS— resultaba además extraordinariamente atractiva para el equipo de investigación, porque abría interesantísimas posibilidades para el futuro. A medio plazo, y de forma muy destacada, permite la interoperabilidad con otros sistemas de información relacionados, como bases de datos jurídicas, sitios web de otros estatutos o instituciones, etc., bien sea mediante el mapeado directo o bien a través de sistemas de organización del conocimiento ampliamente aceptados en el mundo jurídico —como, por ejemplo, Eurovoc— que puedan servir de lenguajes de interconexión. Más a largo plazo, se apuntó la complejización de la ontología de manera que pueda servir —junto a un sistema de reglas— como sistema experto para la administración estructurada del corpus jurídico aragonés.

## **4 La ontología como estructura básica de la arquitectura de información del web**

### **4.1 El Estatuto como eje de los contenidos del web: una ontología empírica al servicio de los objetivos del sistema**

La característica más notable de la ontología del Estatuto de Autonomía de Aragón es que no fuerza el conjunto de conceptos y términos del Estatuto dentro de un esquema conceptual ajeno, sino que —aun intentado obedecer a las mejores prácticas en facetación de estructuras conceptuales y, en particular, teniendo en cuenta la arquitectura de otros sistemas de organización del conocimiento en el

campo del Derecho— surge del propio estatuto de forma natural. Esto es así porque el propio Estatuto constituye, como se ha explicado, el centro y el eje del sitio web, y, por tanto, del sistema de navegación.

La ontología controla la categorización de todos los recursos del sitio web y los pone en relación con el estatuto. Todos los recursos son etiquetados con términos preferentes de la ontología, de manera que, para cada elemento del texto del Estatuto —título, capítulo, artículo— se muestran todos los recursos de información relevantes para los conceptos contenidos en él.

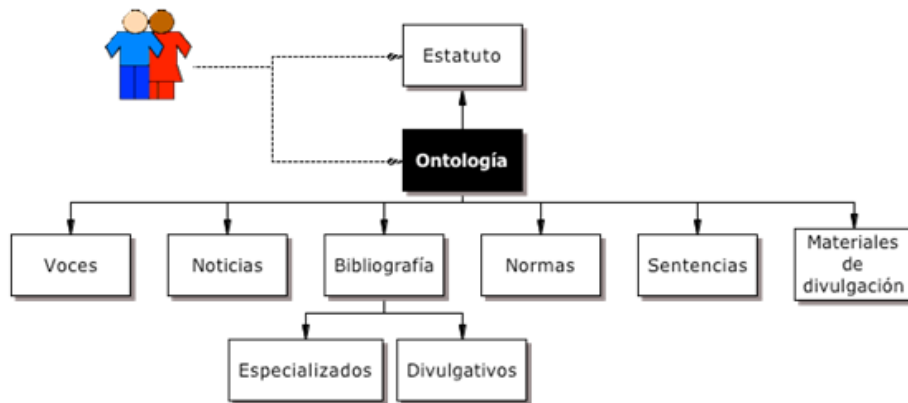


Fig. 2. La ontología como interfaz entre el Estatuto y los recursos.

## 4.2 Elaboración de la ontología

La ontología se elaboró de forma inductiva, partiendo de todas las palabras clave del Estatuto, que se agruparon primero en términos, posteriormente en conceptos y finalmente, se estructuraron jerárquica y asociativamente para formar un sistema terminológico y, desde el punto de vista de la organización del conocimiento, un vocabulario estructurado.

Consta de 1134 conceptos con sus sinónimos y términos relacionados, y se estructura en siete árboles: régimen político, actividad económica y social, personas y colectivos, instituciones del Estatuto, administraciones y servicios públicos (actividad pública), políticas públicas y normativa (Figura 3). También se han explicitado 52 desambiguaciones.



Fig. 3. Arquitectura básica de la ontología.

## 4.3 El isomorfismo entre el Estatuto y la ontología y el etiquetado del Estatuto

La estructura del Estatuto en títulos, capítulos y artículos constituye un esquema conceptual en sí mismo y, lógicamente, el mecanismo de navegación por excelencia



del mismo. Sin embargo, la estructura textual del Estatuto responde a una organización que sirve lógicamente a los intereses y los usos idiosincrásicos de los agentes implicados en el dominio específico de los estatutos de autonomía, y, concretamente, de los estatutos de autonomía españoles.

Por ello, la ontología intenta proporcionar una perspectiva más universal que puede ayudar a recuperar la información desde puntos de vista diferentes. Básicamente, el modelo conceptualiza el Estatuto como una herramienta jurídica que, respetando y postulando unos principios jurídicos, regula la actividad política y social aragonesa en beneficio de los ciudadanos estructurando políticas públicas que se desarrollan mediante instituciones de acuerdo a un conjunto de normas superiores o dependientes (Figura 4).

Sin embargo, y además de proporcionar una estructura conceptual alternativa, la ontología proporciona fundamentalmente una ligazón entre los términos del Estatuto, suficientemente ricos y variados en sí mismos para servir como puerta de acceso tanto a técnicos como a ciudadanos en general.

Finalmente, para facilitar el mantenimiento del sistema, se ha desarrollado un procedimiento automático que permite importar y exportar la ontología, y otro que permite anotar y etiquetar automáticamente los términos del Estatuto en XML, ligándolos así a los nodos de la ontología.

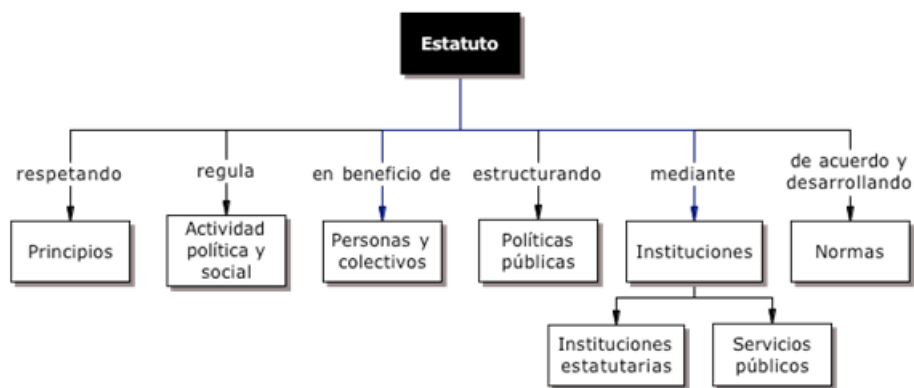


Fig. 4. El modelo conceptual subyacente.

## 5 Implementación de la web semántica

### 5.1 Ontología

En el desarrollo del web del Estatuto de Aragón se hace uso de los estándares citados en los apartados anteriores para complementar los contenidos expuestos en el web con "información semántica". Como hemos dicho se emplea un sistema de categorización conceptual conforme con la recomendación Simple Knowledge Organization System (SKOS). El empleo de este sistema conceptual permite que los investigadores y administradores del web comprendan fácilmente la organización de los conceptos y sus relaciones.

Para su implementación, se ha definido una sintaxis sencilla que permite definir dicho árbol conceptual mediante un fichero de texto sencillo. Dentro del mismo, por medio de un conjunto de símbolos se definen visualmente las relaciones entre los conceptos. Dicho fichero es importado y procesado automáticamente en el sistema. Gracias a esto, se permite que la ontología conceptual del web pueda ser mantenida y mejorada por los colaboradores, que no son especialistas en informática.

La información anterior es procesada internamente mediante sistemas de indexación y cacheo para ser empleada en los algoritmos internos de búsqueda y relación de contenidos del web. Adicionalmente se incorpora un módulo externo para su publicación en formato RDF, lo que permite ponerla a disposición de otros webs que en el futuro pudieran aprovecharla.

## 5.2 Arquitectura lógica

En el siguiente diagrama puede observarse la estructura de componentes y procesos aplicados para dar solución a los requerimientos planteados en el proyecto. Se describe toda la cadena de procesos desde que se carga la ontología y se procesa, hasta su empleo dentro de las búsquedas del web y su publicación mediante RDF.

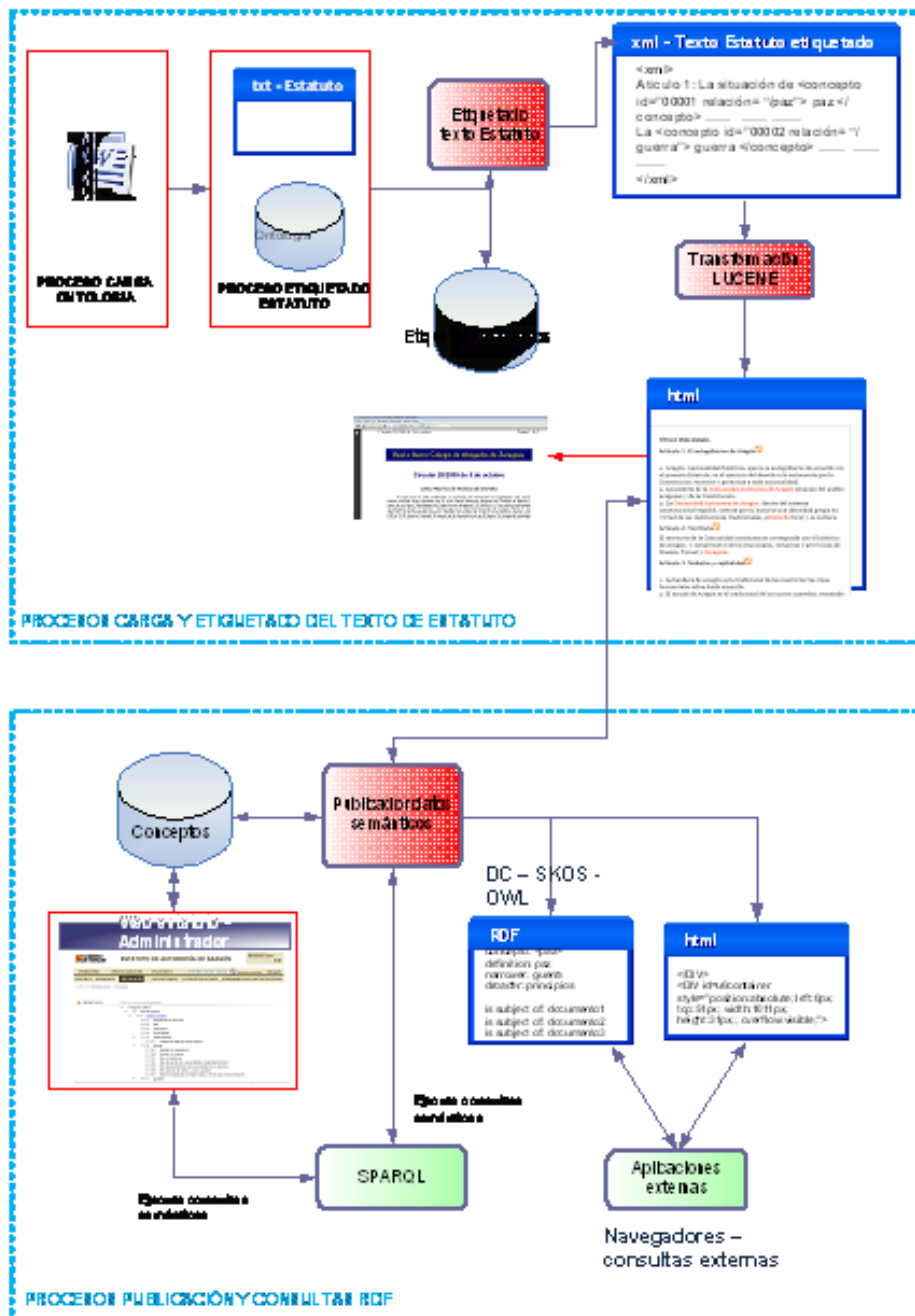


Fig. 5. Arquitectura lógica

## 5.3 Subsistemas

### 5.3.1 Subsistema de carga de la ontología y etiquetado del texto del estatuto

El sistema se divide en dos subsistemas principales. El primero de ellos carga en el sistema la ontología desde el fichero de texto en el que ha sido definida y, a continuación, etiqueta el texto del estatuto con el dialecto XML definido al efecto. Este subsistema está compuesto por los siguientes procedimientos automatizados:

- Proceso de carga ontología: Proceso de carga automática de la ontología partir del mencionado documento de texto. Tras el procesado dicho árbol conceptual es almacenado en un modelo de datos interno optimizado para lograr un mejor rendimiento.
- Proceso de etiquetado del texto del Estatuto: Mediante el cual el sistema toma como entrada dichos conceptos de la ontología y el texto del Estatuto. Realizando un recorrido sobre ambos detecta la ocurrencia de términos conceptuales dentro del texto del Estatuto y crea simultáneamente las relaciones entre conceptos y palabras asociadas. El resultado es almacenado en un fichero XML que contiene la versión procesada y etiquetada del texto.

A partir del fichero XML resultante de los dos procesos anteriores, mediante un conjunto de operaciones de transformación se genera el HTML presentado al usuario en la Zona de lectura del Estatuto. El fichero XML contiene las etiquetas necesarias para asociar las palabras y expresiones del texto del Estatuto con las categorías de la ontología y, en base a estas, con los recursos del web.

### 5.3.2 Subsistema de publicación y consulta

El sistema también administra el etiquetado de cualquier documento o contenido publicado en el web mediante los términos de la ontología pertinentes. Así, junto con la información generada en los procesos anteriores, se dispone de una relación completa entre todos los contenidos y texto del Estatuto. Esta información alimenta el servicio de consultas y búsquedas dentro de la propia web, ya sea como búsquedas realizadas explícitamente por el usuario o como búsquedas implícitas realizadas para sugerir al usuario contenidos relacionados en función de los apartados del texto que esté leyendo en ese momento.

Por otra parte, toda la información disponible en el web es publicada conforme al estándar RDF. Para cada documento o contenido se publica su versión en RDF, identificada mediante una URI única, incorporando propiedades del formato Dublin Core para su catalogación y complementándolas con la representación SKOS de las categorías de la ontología.

## 6 Conclusion

El proyecto del web del Estatuto de Aragón ha permitido experimentar sobre el desarrollo de nuevas herramientas para el acceso a textos jurídicos y demostrar su virtualidad, tanto en la utilización por usuarios expertos, fundamentalmente de formación jurídica, como por personas simplemente interesadas en conocer mejor el Estatuto o algunos de los aspectos relacionados con el mismo.

El ámbito acotado de web, al estar centrado en una única norma, ha hecho posible desarrollar con un esfuerzo limitado dos herramientas de web semántica (la ontología y el etiquetado XLM de texto del Estatuto) y utilizarlas en un interface innovador que pone en relación el propio texto de la norma con la ontología y, a través de ella, con los recursos incorporados por el equipo de colaboradores como contenidos del web.

Las categorías de la ontología permiten superar la estructura sistemática tradicional y sirven de base a una nueva forma de “pensar” el Estatuto. La herramienta desarrollada muestra de forma sencilla e inmediata las distintas facetas de las categorías relacionadas con el Estatuto, ya que muestra de forma sintética tanto las partes del texto del Estatuto que regulan la categoría, como los distintos recursos que existen en relación con la misma que no son sólo técnicos y académicos, sino que comprenden incluso noticias que dan cuenta de cómo esa categoría se plasma en la realidad política y social.

## **Referencias**

1. Järvenpää, M., Virtanen, M., Salminen, A.: Semantic Portal for Legislative Information. In: Wimmer, M.A., Scholl, H.J., Grönlund, A., Andersen, K.V. (eds.) Electronic Government EGOV2010. LNCS, vol. 4084, pp. 219--230. Springer, Heidelberg (2006)
2. ISO/DIS 25964-1. Information and documentation — Thesauri and interoperability with other vocabularies — Part 1: Thesauri for information retrieval. Geneve: ISO, 2009.
3. ISO/DIS 25964-2. Information and documentation — Thesauri and interoperability with other vocabularies — Part 2: Interoperability with other vocabularies. Geneve: ISO, 2010.
4. World Wide Web Consortium. SKOS Simple Knowledge Organization System Reference. W3C W3C Proposed Recommendation 15 June 2009. [Latest version available at <http://www.w3.org/TR/skos-reference>]