



ESTUDIOS Y ARTÍCULOS

Datos abiertos entrelazados para una mejor información jurídica: estado actual del Identificador Europeo de Legislación en España

Pedro Juan Baquero Pérez¹

*Universidad de La Laguna
España*

ORCID: [0000-0002-5545-0450](https://orcid.org/0000-0002-5545-0450)

RECIBIDO: 2 de octubre de 2020

ACEPTADO: 9 de marzo de 2021

RESUMEN: La Unión Europea ha promovido la iniciativa European Legislation Identifier (ELI). En esta iniciativa se desarrolla la especificación técnica ELI, la cual define un marco flexible, de forma que cada país se adapte a sus peculiaridades con el fin de mejorar la información jurídica y tratando la legislación como datos abiertos y entrelazados. Su implementación se basa en el concepto de la WEB Semántica, y trata de facilitar el acceso, la interoperabilidad y reutilización de la información de los sistemas de publicación de normativa legal europeos, nacionales y regionales. En este trabajo, basado en la especificación ELI, se ha desarrollado un Marco Teórico que describe un modelo de evolución, tanto tecnológico como de servicios, para la aplicación de los datos abiertos entrelazados a la publicación de la legislación. En base a este Marco Teórico, se ha analizado el estado actual de la implantación en España de ELI.

PALABRAS CLAVE: Datos abiertos entrelazados; WEB Semántica; Ontologías Legales; Textos legislativos; Identificador Europeo de Legislación (ELI).

¹ Doctor en Física e Informática por la Universidad de La Laguna. Ingeniero de Telecomunicación por la Universidad Politécnica de Madrid. Profesor asociado en Ingeniería Telemática en la Universidad de La Laguna y Jefe de Servicio de Informática y Comunicaciones en el Gobierno de Canarias.

CONTENIDOS: 1. Introducción. – 2. Objetivos. – 3. Estado de la cuestión. – 4. Marco Teórico. – 5. Metodología. – 6. Resultados y discusión. – 7. Conclusiones. – Bibliografía.

Linked Open Data for better legal information: current status of the European Legislation Identifier in Spain

ABSTRACT: The European Union has promoted the European Legislation Identifier (ELI) initiative. This initiative develops the ELI technical specification, which defines a flexible framework, so that each country can adapt it to its peculiarities, in order to improve legal information and treat legislation as Linked Open Data. Its implementation is based on the Semantic Web concept, and seeks to facilitate access, interoperability and reuse of information from European, national and regional legal publication systems. In this work, based on the ELI specification, a Theoretical Framework has been developed that describes the application of Linked Open Data from a model of technological and service evolution for the publication of legislation. Based on this Theoretical Framework, the current state of implementation of ELI in Spain has been analysed.

KEYWORDS: Linked Open Data; Semantic Web; Legal Ontologies; Legal Texts; European Legislation Identifier (ELI).

1. Introducción

En un Estado de Derecho la legalidad es uno de sus principios fundamentales, lo que obliga a dar público conocimiento de la normativa a los ciudadanos a través de su publicación (Rodríguez-Vela, 2011). Es la publicidad mediante su publicación en un medio oficial lo que constituye una garantía básica para su entrada en vigor, es decir, para su aplicación dentro de la sociedad. Esto se concreta en los artículos 129.1 y 131 de la *Ley 39/2015, de 1 de octubre, del Procedimiento Administrativo Común de las Administraciones Públicas*, donde, con el fin de una buena aplicación del principio de transparencia, "*las Administraciones Públicas posibilitarán el acceso sencillo, universal y actualizado a la normativa en vigor y los documentos propios de su proceso de elaboración*" y, además, se establece que las Administraciones Públicas podrán utilizar otros medios de publicidad complementarios. Esto último es relevante, ya que promueve el uso de otros medios diferentes a los boletines oficiales para ofrecer una mejor información jurídica, dado que, aunque la publicidad oficial juega un papel fundamental como garantía jurídica, los defectos relacionados con la carencia de sistematizar las sucesivas modificaciones sufridas a través de normativas posteriores, llevaría aparejado una dificultad para el conocimiento y publicidad jurídica a la sociedad. No que hay que olvidar que el artículo 6.1 del Código Civil recoge que el desconocimiento o la ignorancia de la existencia de la norma no puede ser una excusa para su no cumplimiento. En este sentido, aunque esto no implica que toda la sociedad conozca toda la normativa, ya que esto es imposible, sí supone que los poderes públicos deberían publicitar las normas como garantía más allá de únicamente su publicación oficial.

Por tanto, se puede afirmar que con la sola publicación en abierto de los boletines o diarios oficiales no se cumple con los principios de mayor transparencia. En España, la regulación de la transparencia se marca en la *Ley 19/2014, del 29 de diciembre, de transparencia, acceso a la información pública y buen gobierno* y en cada una de las leyes de las Comunidades Autónomas. En estas leyes las normativas legales son tratadas entre los objetos que merecen una mención especial, y se establece que las Administraciones Públicas deben acometer la revisión, simplificación y, cuando proceda, la consolidación normativa, con el fin de dar mayor seguridad jurídica para que la normativa en vigor no esté dispersa en distintas normas. Por otra parte, el artículo 7 de la norma estatal establece sobre qué documentos deben darse publicidad en los pasos previos a la publicación de las normas legales, para que sean sometidos a información pública en su tramitación, como son los dictámenes, anteproyectos de Ley y memorias de los expedientes preparatorios.

Además, no es solo un objetivo que estos documentos estén disponibles de forma libre por el público en general, sino también que esta normativa pueda ser reutilizada y referenciada de la forma más efectiva posible. En España, la regulación de la reutilización de la información se encuentra en la *Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público*. Esta Ley establece un marco mínimo, donde se recoge las condiciones que los distintos documentos generados por el sector público sean susceptibles de reutilización. Los diarios o boletines oficiales son documentos generados por las Administraciones Públicas, con lo que entran dentro de esta Ley.

El desarrollo de sistemas de información que gestionan el contenido, los datos y los metadatos de los textos jurídicos ha sido una práctica común con el fin de que la legislación sea fácilmente accesible al público. Casanovas *et al.* (2016) pone ejemplos de distintas publicaciones a nivel internacional y cómo estas han ido incorporando la publicación electrónica de la legislación con formatos estructurados. También Alvite Díez (2009) hace un recorrido sobre el uso de formatos estructurados para los textos legales en España, donde a principios de este siglo ya existían iniciativas que disponían de publicaciones en formatos estructurados usando principalmente *eXtensible Markup Language* (XML); aunque todavía existen sistemas que no ofrecen formatos estructurados en sus textos publicados electrónicamente.

En la Unión Europea se han realizado esfuerzos incorporando datos estructurados XML en los recursos legales (Rodríguez-Doncel *et al.*, 2018). En cualquier caso, sigue siendo uno de los desafíos en los últimos años el mejorar la interoperabilidad entre los distintos sistemas de publicaciones legislativas y facilitar la reutilización de esta información. Entre los problemas existentes tenemos la diversidad de fuentes, ya que aunque muchos disponen de documentos estructurados, cada uno describe sus datos de forma diferente. Esto dificulta el acceso, la reutilización y la interoperabilidad de la información jurídica. Esta diversidad de las fuentes ha dificultado que se dispusiese de un punto único donde se pueda acceder de forma ágil y eficiente a toda esta normativa distribuida por diversos sistemas. En este sentido, la Unión Europea ha creado la iniciativa *European Legislation Identifier* (ELI), definida en las *Conclusiones del Consejo en las que se promueve la creación de un Identificador Europeo de Legislación (ELI)*, que ha sido adoptada en el año 2012 por los países de la Unión Europea, con el fin de permitir el acceso a la normativa europea en un formato normalizado. A partir de la experiencia de las implementaciones realizadas, esas Conclusiones fueron actualizadas en 2017 a través de las *Conclusiones del Consejo de 6 de noviembre de 2017 sobre el Identificador Europeo de Legislación*.

En este sentido, es necesario que los sistemas responsables de las publicaciones oficiales se unan a esta red de intercambio de textos legislativos y comprendan este marco tecnológico y ontológico. El marco que ha definido la Unión Europea, a través del proyecto ELI, es lo suficientemente flexible para que cada país se adapte

a sus peculiaridades. Esto supone que se tienen altos grados de libertad, pero siempre cumpliendo unos requisitos mínimos. También, se requiere que cada país de la Unión Europea defina su propia estrategia de implantación de ELI adaptada a sus peculiaridades.

En el caso de España se ha definido la Fase 1 de la especificación técnica española ELI², que se ha adaptado a su configuración territorial, donde el Estado y las Comunidades Autónomas disponen cada uno de sus propias unidades de publicación. Una característica de esta especificación española, en línea con la recomendación europea ELI, es también su flexibilidad, de forma que cada Comunidad Autónoma pueda adaptarse a sus sistemas. El proyecto ELI español está en fase de implantación, existiendo, a parte del Estado, algunas Comunidades Autónomas que ya han implantado su primera fase, mientras su segunda fase todavía no ha sido diseñada. Los sistemas que lo han implantado lo han hecho de forma desigual, siguiendo distintas estrategias. En cualquier caso, ELI no es requisito necesario y suficiente para mejorar la calidad del acceso, reutilización, e interoperabilidad y, por tanto, mejorar la información jurídica a la sociedad, aunque sí se puede presentar como su soporte fundamental.

Este trabajo se estructura de la siguiente manera: Primero, presentamos los objetivos del trabajo. En segundo lugar, se hace un recorrido por el estado de la cuestión, donde se trata la situación general de la implantación de ELI en Europa y España; los servicios que se pueden ofrecer soportados en ELI; y las peculiaridades del caso español, así como qué servicios básicos se necesitan ofrecer para cumplir con los principios de transparencia. En tercer lugar, se desarrolla un Marco Teórico que contempla los servicios y tecnologías de un sistema general de información jurídica para España. A continuación, basados en este Marco Teórico, se establece una metodología para medir el estado actual en España, para a continuación, presentar los resultados y realizar una discusión de los mismos.

2. Objetivos

El objetivo principal de este trabajo es conocer el estado actual de los sistemas de información jurídica soportados en la especificación ELI europea en España. Para lo que se ha planteado la pregunta de investigación: *¿En qué medida los sistemas actuales en España tienen implementados los servicios y tecnologías necesarias para ofrecer una mejor información jurídica a la sociedad basados en el Identificador Europeo de Legislación?*

Para contestar a esta pregunta, se ha planteado que debe existir un Marco Teórico que sea el punto de referencia para disponer de una visión ideal y a largo plazo de lo que debería contener una solución conceptual de los sistemas que garantizan

² Especificación técnica para la implementación del Identificador Europeo de Legislación en España. (Fase 1). Esta especificación se encuentra publicada en <https://www.elidata.es/>. Consultado el 5 de febrero de 2020.

una mejor información jurídica. Este marco debe contener una visión técnica, relacionada con las tecnologías, y una visión estratégica, relacionada con los servicios que se pueden prestar.

La visión técnica debe basarse no solo en la especificación ELI, sino también soportada en el estado del arte de las tecnologías de los datos abiertos entrelazados y WEB Semántica, y contener los aspectos semánticos legales. La visión estratégica nos da las respuestas a las potencialidades que provee la base técnica ELI, en el sentido de que los sistemas ofrezcan servicios de valor añadido relacionados con el acceso, la reutilización y la interoperabilidad aplicando ELI y las tecnologías asociadas. Por tanto, otro objetivo de este trabajo es obtener un Marco Teórico de los servicios y tecnologías de un sistema general de información jurídica para España.

3. Estado de la cuestión

Como se ha visto, por parte de la Unión Europea (UE) ha existido una preocupación en proporcionar normas jurídicas de manera abierta, accesible e interoperable, para lo que ha definido la especificación ELI. En este sentido, se han hecho esfuerzos por adoptar tecnologías de la WEB semántica y de datos abiertos entrelazados para que el contenido legislativo sea más fácilmente accesible al público. En este apartado, se describirá la especificación ELI, las potencialidades de ELI para prestar servicios, el estado actual en Europa y en España. Previamente, se hará una descripción básica de las tecnologías de la WEB semántica y de datos abiertos entrelazados.

3.1. Conceptos de las tecnologías de la WEB semántica y de datos abiertos entrelazados

Los conceptos WEB semántica y de datos abiertos entrelazados son la base tecnológica de ELI, y están altamente relacionados. Cuando hablamos de datos abiertos no solo entendemos los datos como tales, sino que también se incluyen dentro de la misma definición a los documentos. Estos dos conceptos se sustentan en las siguientes bases (Berners-Lee, Hendler y Lassila (2001)): se debe permitir que los sistemas de información puedan libremente gestionar la combinación y la integración de los datos, de forma que cualquiera pueda dar más valor a los datos originales; y los datos se deben almacenar en un lenguaje que su semántica (significado) sea almacenado de forma que los sistemas de información (y también las personas) puedan entenderse entre sí.

Actualmente, muchas administraciones están cada vez más inmersas en incluir entre sus prioridades el concepto de datos abiertos, de forma que los datos que se generan en las administraciones estén disponibles por las personas sin limitaciones de forma, lo cual facilita la reutilización para cualquier propósito. Junto con este concepto de datos abiertos (Open Data) está el término Linked Data (Datos Entrelazados). Estos dos conceptos no son del todo equivalentes: Linked Data se

refiere a datos que están relacionados a través de tecnologías que aporta la Web Semántica, principalmente usando RDF (Resource Description Framework); y Open Data se refiere, además, que sean abiertos, es decir, gratuitos y libres de uso. Si unimos estos dos conceptos hablamos de Datos Abiertos Entrelazados (LOD).

Existe un conjunto amplio de estándares tecnológicos que sustentan los servicios de la WEB semántica. Entre estas tecnologías destacamos XML, RDF, *Web Ontology Language (OWL)* y *SPARQL Protocol and RDF Query Language (SPARQL)*³, y adoptando vocabularios y ontologías específicas en los documentos legislativos.

XML es un lenguaje promovido por el World Wide Web Consortium (W3C) que permite a las máquinas la descripción y la lectura de datos. Se trata de un estándar abierto que está ampliamente extendido. En un fichero XML los datos se encuentran delimitados por etiquetas que dan información sobre ellos y, al mismo tiempo, los estructuran.

Para la utilización de documentos XML en los navegadores se utiliza eXtensible HyperText Markup Language (XHTML). XHTML es un archivo HTML formalizado a través de un XML válido, es decir, es una versión de XML en HTML. Realmente, XHTML usa una sintaxis muy similar a HTML, pero es más estricta en el marcado, y todos los elementos tienen que estar etiquetados para evitar ambigüedades.

Mientras XML facilita la descripción de un documento, RDF está pensado para el intercambio de datos. Con RDF se establecen relaciones entre distintos documentos que pueden estar alojados en el mismo sitio o a lo largo de toda la WEB. RDF se considera como uno de los pilares de la WEB Semántica. Así, se puede expresar cualquier hecho de forma estructurada de forma que los sistemas de información puedan tratarlos para generar más valor. Y a través de las distintas relaciones con significado se pueden generar redes de conocimiento que describen las relaciones entre los distintos documentos.

En resumen, la mayor diferencia entre RDF y XML es que RDF está concebido para la representación distribuida del conocimiento y está enfocado mayormente al significado. RDF no tiene que usar el formato XML, ya que también usa metadatos para describir los datos de los documentos. En cualquier caso, se puede incrustar RDF en XML usando la notación XML.

Además, existen otras tecnologías que se están extendiendo cada vez más, como la denominada JavaScript Object Notation for Linked Data (JSON-LD). Esta tecnología usa los formatos JSON, el cual es otro formato para el intercambio de datos y se basa en el lenguaje conocido de programación JavaScript. Se trata de una alternativa a XML que se ha adoptado ampliamente y actualmente se

³ En realidad existe un amplio conjunto de tecnologías asociadas a la Web Semántica. Para mayor información sobre todas las tecnologías nombradas en este artículo se puede consultar el World Wide Web Consortium (W3C) <https://www.w3.org/>, donde continuamente se actualizan los distintos estándares de estas tecnologías, de forma que se asegure la evolución de la Web a largo plazo. .

considera independiente del lenguaje JavaScript. Entre sus ventajas es que es más sencillo de usar por un analizador sintáctico. También, existen otros formatos diferentes de RDF como *Terse RDF Triple Language* (TURTLE), el cual es otra alternativa a RDF/XML, que no se basa en XML, y que gestiona mejor las redes de conocimiento válidos RDF. Se considera que tiene mejor lectura y edición que XML.

Para darle más significado, cuando se interpretan ficheros a lo largo de una red de relaciones, es necesario el uso de vocabularios. Esta es una de las funciones principales de OWL, el cual es un lenguaje enfocado en el desarrollo de ontologías y la descripción de taxonomías. Una ontología es una especificación explícita y formal de una conceptualización que es compartida por distintos agentes, por ejemplo, agentes que usan un vocabulario y lenguaje específico del ámbito jurídico. Esta conceptualización es un modelo abstracto de la realidad a la que representa. De este modo, una ontología se puede ver como una forma esquemática de crear un lenguaje común, con sus reglas sintácticas y, así, se facilita que distintos sistemas hablen y compartan información entre sí.

Por otra parte, una taxonomía es una ordenación jerárquica y sistematizada de grupos. Los grupos pueden ser de cualquier tipo y dentro de un ámbito. Se suele asociar a un conjunto de términos consensuados en un ámbito determinado. Es jerárquico, dado que los términos no se clasifican en una estructura lineal, sino que dentro de un término se pueden definir otros que los categorizan con más detalle. Por ejemplo, si se define una materia como *Educación*, se pueden definir otras submaterias dentro de esta, como *Universitario* y *No Universitario*.

En definitiva, el vocabulario adicional que se incluye en una ontología permite que los ficheros XML y RDF se estructuren semánticamente de forma adecuada, ayudando a establecer una más rica correspondencia entre estos ficheros.

Por último, la aportación de ficheros que puedan ser accedidos por máquinas es el primer paso para garantizar el acceso a los datos abiertos. En este sentido, la principal tecnología de la WEB semántica es SPARQL, el cual es un lenguaje para la consulta semántica sobre RDF. SPARQL está estandarizado por W3C y permite hacer preguntas a distintos objetos escritos en RDF con el fin de obtener información más rica. Dado que RDF establece relaciones entre documentos distribuidos en una red, las preguntas que se hacen con SPARQL se responde recorriendo todos estos documentos, con lo que se pueden responder preguntas complejas.

Las consultas SPARQL, enfocadas al descubrimiento de información en las redes de conocimiento suelen ser complejas, por lo que se requiere un conocimiento especializado a los programadores. Por ello, la mayor parte de los portales ofrecen servicios como el *Representational State Transfer* (REST) que resultan menos complejos, aunque más limitados.

3.2. La especificación ELI

Con la introducción de ELI, en 2012, por el Consejo Europeo, esta especificación se ha convertido en el nuevo marco común que deben adoptar los sistemas nacionales de publicación jurídica. ELI se basa en una *Uniform Resource Identifier* (URI) normalizada que identifica los recursos jurídicos en la WEB. El término en español de URI es un *Identificador Uniforme de Recursos*, el cual permite acceder a un recurso en Internet. Así, una URI permite identificar una página WEB o un documento alojado en una WEB, es decir, está fuertemente relacionado con las conocidas URLs.

Además, ELI proporciona una ontología OWL, que expresa los metadatos de los textos jurídicos. ELI no es un modelo único, que puede y, en muchos casos, debe ampliarse para adaptarse a las especificidades de los diferentes sistemas de los Estados miembros de la UE. También, con la misma visión, se desarrolló el identificador de jurisprudencia europea (ECLI) (van Opijnen e Ivantchev, 2015), sin conexión con el ELI, pero con el objetivo de gestionar la jurisprudencia. Por lo tanto, ELI y ECLI son iniciativas independientes, pero tienen los mismos principios. Este trabajo se centra en ELI.

El objetivo es tener una red de conocimiento sobre legislación, donde la información y el conocimiento encuentren un canal para que las unidades encargadas de publicar los textos legales en los diferentes países tengan interoperabilidad entre sus sistemas de información, rompiendo no sólo las barreras debidas a la existencia de diferentes sistemas legales sino también las soluciones técnicas que cada unidad ha implementado. Entre algunas ventajas de la especificación ELI tenemos la mejora de la interoperabilidad, el acceso a la normativa, la incorporación de nuevos servicios y la transparencia, al tiempo que se mejoran los costes.

El esquema principal de ELI se basa en dos aspectos. En primer lugar, la creación de identificadores URIs para la legislación, de forma que todos los textos se identifican de la misma manera. En segundo lugar, la misma descripción de sus metadatos mediante una ontología basada en los *Requisitos Funcionales de los Registros Bibliográficos* (FRBR)⁴, la cual se trata de una ontología específica para modelar conceptualmente los registros bibliográficos. ELI utiliza FRBR utilizando tecnologías de datos abiertos entrelazados. Por lo tanto, es capaz de representar las relaciones entre los tipos de recursos legislativos y desde diferentes puntos de vista. Además, asocia características, como las fechas, que son necesarias para marcar un recurso legal, y vínculos para relacionar los actos legislativos. De esta

4 Grupo de Estudio IFLA sobre FRBM (2009). Functional Requirements for Bibliographic Records. Final Report. Technical report, International Federation of Library Associations and Institutions. https://www.ifla.org/files/assets/cataloguing/frbr/frbr_2008.pdf. Consultado el 5 de marzo de 2020.

manera, los metadatos jurídicos definen una red mundial de entidades interconectadas, y ELI facilita la difusión de los metadatos estructurados de todos los sistemas que utilizan este modelo.

La especificación ELI se apoya en tres pilares⁵ asociados, respectivamente, a los identificadores URI, los metadatos y un lenguaje para el intercambio de datos. Estos pilares pueden aplicarse independientemente unos de otros. Sin embargo, cuando se combinan, surgen las ventajas y las potencialidades de ELI.

El pilar 1 utiliza URIs sobre HTTP para identificar los textos jurídicos oficiales publicados electrónicamente. Estos identificadores se hacen a través de plantillas de URI ELI que pueden ser tratadas automáticamente. Esta URI se basa en la aplicación de componentes con semántica tanto legal como de usuario final. Cada país puede diseñar sus propias URIs que son autodescriptivas, pero tratando de utilizar, siempre que sea posible, los componentes que han sido descritos en la especificación ELI. En otras palabras, cada país es libre de seleccionar estos componentes y de organizarlos de la mejor manera posible para satisfacer sus necesidades. Estos componentes se han definido con la siguiente estructura:

```
/eli/{jurisdiction}/{agent}/{sub-agent}/{year}/{month}/{day}/{type}/{natural_identifier}/{level1} . . . /{point in time}/  
{version}/{language}6.
```

El pilar 2 desarrolla los metadatos, que se describen mediante una ontología que formaliza un conjunto de conceptos y relaciones en el ámbito de las publicaciones legislativas. Esto permite un entendimiento común, evitando ambigüedades entre los términos que utiliza cada sistema. Al estar formalizado, es decir, que utilizan los mismos metadatos, permite el procesamiento automático a través de máquinas. Existen un conjunto muy amplio de metadatos. Un ejemplo de metadato es *title* para indicar el título del documento. En cualquier caso, cada país es libre de aplicar sus propios esquemas para sus metadatos, aunque se fomenta que se siga la especificación ELI. En este sentido, los metadatos ELI se formalizan a través de la ontología ELI. El grupo de trabajo europeo ELI está a cargo de la actualización y el mantenimiento de la ontología. Para ello se ha utilizado el modelo FRBR, que se adapta a otras iniciativas estandarizadas dentro del campo de la información jurídica. Para apoyar esta ontología, se aplica ELI/XML que, como se ha visto, es una codificación de los metadatos ELI en un XML. ELI/XML puede utilizarse independientemente o incrustarse en otros documentos XML.

5 Publications Office of the UE. <https://op.europa.eu/es/web/eu-vocabularies/eli>. Consultado el 13 de marzo de 2020.

6 Dos ejemplos de cómo se han aplicado son: en Francia, `eli/{type}/{year}/{month}/{day}/{natural_identifier}/{version}/{level1}`; y en Serbia, `/eli/{level1}/{type_gazette}/{passed_by}/{type_document}/{year}/ {number}/ {serial_number}/`

Por último, el pilar 3 está diseñado para permitir una distribución e intercambio de datos eficiente. Los archivos de intercambio deben tener el mismo lenguaje o formato. Para ello, se ha optado por la recomendación del W3C para describir los metadatos utilizando el *Resource Description Framework in Attribute* (RDFa), que define la incorporación de RDF en XHTML. Además, se promueve el uso de otras tecnologías como JSON-LD y TURTLE. En cualquier caso, también, cada país es libre de añadir otros formatos.

En España, el análisis de la implantación del ELI ha sido realizado por la Comisión Sectorial de Administración Electrónica, integrada por el Gobierno español, las Comunidades Autónomas y las Autoridades Locales. Esta Comisión creó el grupo de trabajo ELI en 2017. El resultado de su trabajo fue la adopción de la *Especificación Técnica para la Implementación del Identificador Europeo de Legislación en España (Fase 1)* en marzo de 2018, la cual es una versión abreviada de la europea. Esta especificación tiene como objetivo establecer las directrices para la implantación de ELI en España siguiendo las recomendaciones europeas. En esta especificación se ha seguido la estrategia europea flexible y gradual. En este sentido, la implantación se ha propuesto en diferentes fases, empezando por los textos españoles y regionales, dejando para otra fase la incorporación de los boletines provinciales. Esta implantación se define a partir de unos requisitos mínimos, donde cada sistema puede adaptarse a sus peculiaridades. La fecha límite establecida para su finalización era diciembre de 2019.

En cuanto al identificador ELI (Pilar 1), la especificación española simplifica la recomendación europea para adaptarse mejor a las peculiaridades de España. Así, contiene el identificador *language*, permitiendo una interconexión eficiente de la legislación en entornos jurídicos multilingües y regionales: la jerarquía de los FRBR permite tener un recurso jurídico descrito sólo una vez asociado a múltiples traducciones. Por otro lado, cada Comunidad Autónoma tiene su propio identificador, denominado *jurisdiction*, lo que significa que pueden interrelacionar los recursos publicados en los diferentes sistemas. Entre los identificadores que no se han incorporado se encuentran los correspondientes a las diferentes subdivisiones de los textos; así, no se incorporan los identificadores *level 1, ...* Esto último no facilita la fragmentación de los textos (por ejemplo, dividiendo un texto en título, capítulo, sección, artículo, párrafo) usando ELI, aunque cada sistema puede incorporarlos a su discreción, pero sin tener que mantener una correlación entre ellos, lo que puede dificultar el establecimiento fácil de relaciones entre los diferentes sistemas en el futuro.

Con respecto al Pilar 2, la especificación española ha recogido un subconjunto muy reducido de metadatos de la recomendación europea como obligatorios, el resto son opcionales. En esta especificación, solo es obligatorio un subconjunto de metadatos ELI: *type_document*, *version_date* (necesario cuando se trabaja con recursos consolidados), *format*, *title*, *date_publication* (necesario cuando se utiliza el subtipo de corrección), *realizes*, *is_realized_by*, *embodies*, y *is_embodied_by*. Así, no

ha incorporado como obligatorios metadatos que permitan establecer relaciones con sectores o materias, como los metadatos *relevant_for* y *is_about*. Tampoco ha tenido en cuenta los metadatos que identifican la fuente, como *pass_by* y *transposed_by*. Esto dificulta la trazabilidad de la procedencia del texto, lo cual es fundamental para dar más transparencia a la legislación. Por último, no se incorporan metadatos para dar información sobre su vigencia, como *in_force*, *first_date_entry_in_force*, *date_no_longer_in_force* y *date_applicability*.

Por último, la especificación española, en su Fase 1, no hace obligatoria la incorporación del Pilar 3. Es decir, no incluye los archivos con formato RDFa. Aunque esto no ha supuesto que todas las soluciones que han implementado ELI no lo hayan incorporado a través del formato XHTML.

3.3. Estado actual de la implantación de ELI

En la actualidad hay varios países, incluida la misma unidad europea encargada de las publicaciones, que han iniciado la aplicación de la iniciativa ELI mediante diversas soluciones. Según EUR-Lex⁷, la han aplicado Albania, Austria, Bélgica, Croacia, Dinamarca, Finlandia, Francia, Hungría, Irlanda, Italia, Luxemburgo, Malta, Noruega, Polonia, Portugal, Serbia, España y Reino Unido. y la Oficina de Publicaciones. El ritmo, la estrategia de aplicación y la solución han sido generalmente diferentes entre sí. Así, encontramos países como Portugal, Italia, Dinamarca, Francia, Reino Unido, Irlanda y Noruega que han implementado una solución que incluye los tres pilares, aunque con diferentes alcances en los metadatos elegidos, incluso algunos con sólo algunos metadatos básicos. En general, todos publican con el formato de esquema RDF. Además, cada país se ha adaptado a las especificidades de su propia configuración jurídica. Por último, la Oficina de Publicaciones de la Unión Europea (OPUE), que publica el Diario Oficial de la Unión Europea (DOUE), es la que más se ajusta a la especificación europea ELI.

Como hemos visto, cada país se ha adaptado a sus propias peculiaridades, cada uno seleccionando y organizando los metadatos que mejor se adaptan a sus sistemas. Otros países como Luxemburgo, Finlandia y la OPUE han emprendido soluciones que van más allá del alcance de ELI y han aprovechado el potencial de las tecnologías asociadas a ELI. En el caso de Finlandia⁸, el ELI y el ECLI se han aplicado de forma combinada, de modo que tanto la jurisprudencia como los textos legislativos están relacionados para ofrecer mejores servicios a los

7 ELI-Euro-LEX. URL: <https://eur-lex.europa.eu/eli-register/about.html>. Consultado el 6 de febrero de 2021.

8 Semantic Finlex – Finnish Law and Justice as Linked Open Data. URL:

<https://data.finlex.fi/en/main>. Consultado el 2 de abril de 2020.

diferentes operadores jurídicos. Además, han creado un conjunto de servicios para la consulta de la normativa usando SPARQL y REST⁹.

En el caso de Luxemburgo y la OPUE, han creado redes de conocimiento jurídico dentro de su propio ámbito y las han vinculado entre ambas (Francart *et al.*, 2019). La red de conocimiento jurídico, como fuente de datos para cada canal de difusión, sirve también para difundir conjuntos de datos sobre la legislación en portales de datos abiertos, incluidos los archivos originales, los metadatos de contenido y los vocabularios de apoyo utilizados para indizar la legislación. Entre estas implementaciones se encuentran las realizadas por CELLAR¹⁰ en la OPUE y por Casemates¹¹ en Luxemburgo. En CELLAR el grafo de conocimiento¹² legal se almacena dentro de EUR-Lex. Se trata de un depósito semántico que contiene una descripción estructurada de toda la legislación de la UE (Francesconi *et al.*, 2015). La base de datos de CELLAR puede consultarse libremente desde el exterior utilizando SPARQL. Asimismo, Casemates dentro del portal del Diario Oficial del Gran Ducado de Luxemburgo (Legilux) ha puesto en marcha una red que la vincula con la normativa europea. Estas implementaciones proporcionan un portal WEB único para acceder a todos los contenidos, incluyendo las diversas series de boletines, proyectos legislativos, acuerdos internacionales, tratados y otros más. Además, cuando los usuarios acceden a un texto jurídico de interés, disponer de un grafo de conocimiento les permite navegar por cualquier lugar. De esta manera, pueden buscar legislación con características similares, enumerando textos jurídicos de la misma autoría o con palabras clave de la misma materia.

Por último, en el caso español, a pesar del calendario establecido por la especificación ELI española, que supondría que la Fase 1 debería estar implantada en el año 2020, no todas las Comunidades la han iniciado o completado. Así, en marzo de 2020, la situación de implementación del ELI en España refleja que sólo el Estado, Cataluña, Murcia, País Vasco y Valencia han cumplido con esta Fase 1.

En resumen, ELI ha sido aplicado por los distintos países, cada uno adaptándolo a sus posibilidades. Pero sólo unas pocas iniciativas han aprovechado el potencial de ELI. Se puede decir que por el momento no se puede hablar todavía de que las potencialidades del ELI sean una realidad a nivel transnacional. En el caso español solo cinco sistemas han implementado ELI en una fase inicial.

9 REST API (JSON-LD). URL: <https://data.finlex.fi/fi/rest-api>. Consultado el 2 de abril de 2020.

10 CELLAR. URL: <https://joinup.ec.europa.eu/solution/cellar>. Consultado el 4 de abril de 2020.

11 LEGILUX. URL: <http://legilux.public.lu/editorial/casemates>. Consultado el 16 de abril de 2020.

12 Un grafo de conocimiento es un conocimiento que está organizado de forma que las máquinas puedan fácilmente entenderlo y a la vez extraer información a través de ficheros distribuidos. En un grafo de conocimiento, existen atributos enlazados a través de aristas y vértices.

3.4. Servicios soportados por ELI

Francart *et al.* (2019) afirman que ELI tiene la capacidad de cambiar el mundo jurídico y proporcionar el acceso a la información jurídica no sólo a los usuarios finales sino también a los sistemas de información. En este sentido, plantean una serie de retos y potencialidades que amplían las posibilidades de aplicación de ELI. Entre las grandes potencialidades se encuentra la creación de grafos de conocimiento utilizando las tecnologías LOD (Datos Abiertos Entrelazados). Con el uso de los grafos de conocimiento podemos plantear preguntas generales. Por ejemplo, "¿Qué países han transpuesto una determinada directiva?", "¿Qué traducciones están disponibles para una determinada directiva?", "¿Qué sectores tienen mayores requerimientos regulatorios?" y así sucesivamente. Pero también, preguntas específicas sobre un sector en particular, como: "¿Qué reglamentos existen en toda la UE en relación con la instalación de máquinas recreativas?", "¿Qué países o regiones de la UE tienen más restricciones para su instalación?" y "¿Dónde están las menores cargas fiscales sobre la instalación de máquinas recreativas?". Para lograr que estos grafos de conocimiento puedan responder a este tipo de preguntas es necesario no sólo ampliar la implantación de ELI a un mayor número de países, sino también disponer de más y mejor descripción de la información y de servicios específicos. Este tipo de servicios los denominaremos **servicios en red**, que permiten compartir datos entre la mayoría de los sistemas.

De todos modos, aunque el ELI no ha alcanzado todo su potencial, ya que su uso no se ha extendido, los entornos locales que han desarrollado intensamente ELI, y sus tecnologías asociadas, en sus sistemas de información pueden aprovechar los beneficios que aporta el ELI. Por tanto, antes de contar con un ELI europeo completo, se espera que se pongan en marcha iniciativas en entornos locales o sectores reducidos (como se ha visto en los casos de Luxemburgo y la OPUE) para aprovechar las potencialidades de ELI. Hay que tener en cuenta que ELI facilita en gran medida el acceso a la información jurídica mediante el uso de un mercado legislativo inequívoco. Así, con los metadatos ELI y el uso de formatos abiertos como el RDFa, se está dando más visibilidad a la legislación y se está facilitando el intercambio de datos entre los sistemas de información jurídica dentro del mismo sitio. Francart *et al.* afirman que el aspecto clave que aporta el ELI se puede explotar: está diseñado para trabajar con nuevas arquitecturas de sistemas de información jurídica que se centran en los datos, más que en los documentos. Esto nos permite centrarnos en trabajar con redes de conocimiento jurídico que pueden desarrollarse para prestar servicios en el entorno local. Estos servicios los denominaremos **servicios locales**, que generan servicios dentro de un entorno restringido, aprovechando las características de los datos abiertos entrelazados.

Para el acceso a la información que se ofrece en los servicios locales, se deben prestar servicios como SPARQL y REST, como los servicios prestados por Finlandia, Luxemburgo y la OPUE. A estos servicios los denominaremos **servicios de**

provisión, que facilitan el acceso, la reutilización y la interoperabilidad desde los datos de un entorno local a entornos externos, facilitando el acceso a servicios en red. Además, con ELI, la legislación puede verse mejor en los principales buscadores de Internet. Esto facilita el acceso de los usuarios finales a la legislación, que es uno de los objetivos de ELI. Esto incluye el acceso a la legislación directamente desde la WEB y que los usuarios a través de sus navegadores puedan acceder sin ambigüedades a la legislación que están buscando. Una forma que se está promoviendo para permitir que los motores de búsqueda incorporen información más rica es el uso de schema.org¹³. El cual trata de unificar la forma en que se etiquetan y enriquecen los datos dentro del contenido HTML de una página WEB. De esta manera se mejora la experiencia del usuario para que los motores de búsqueda no sólo la indexen convenientemente, sino que también la clasifiquen y la muestren. En este sentido, para mejorar los resultados de la búsqueda, el grupo de trabajo europeo de ELI ha propuesto introducir dentro de schema.org un vocabulario basado en ELI para la legislación.

ELI puede integrarse con otras ontologías jurídicas, de modo que se establezca un conocimiento más rico. Asimismo, debería hacerse una mayor identificación dentro de los textos legislativos utilizando una mayor fragmentación en los textos jurídicos, lo que permite no sólo crear relaciones a nivel de texto completo, sino también entre niveles más pequeños, como, por ejemplo, entre artículos de diferentes normas (Chalkidis *et al.*, 2017). También, se puede utilizar las características de ELI para prestar otros servicios más sencillos. Por ejemplo, para explotar las posibilidades de ELI de marcar textos legales para crear citaciones legales utilizando el URI ELI. Además, ELI debe desarrollarse o evolucionar para ofrecer servicios de mayor valor. Así pues, debería ampliar su vocabulario o vincularse a un vocabulario jurídico compartido, a fin de facilitar los vínculos entre las reglamentaciones de los diferentes servicios. Por último, ELI puede utilizarse para ampliar su uso a otros ámbitos jurídicos o para ser especificado en un área jurídica específica, no sólo dentro de las publicaciones legislativas, desarrollando la capacidad de la ontología ELI para crear grafos de conocimiento, como el es caso de la experiencia de la utilización de ELI para la normativa específica de protección de datos (Pandit *et al.*, 2018).

En resumen, los tres pilares del ELI no son suficientes para ofrecer una mejor información jurídica a la sociedad, y debe desarrollarse una serie de servicios. Estos servicios los hemos clasificado como servicios en red, servicios locales y servicios de provisión. Además, se puede mejorar los servicios de ELI y aprovechar sus potencialidades para extenderlo a otros ámbitos.

13 Schema.org. URL: <https://schema.org/>. Consultado el 29 de mayo de 2020.

3.5. El caso español

El caso español tiene sus peculiaridades dentro de la UE, lo que hace más importante que en España exista un sistema que facilite el acceso, intercambio y reutilización de las normativas publicadas. Tal y como se ha expresado por parte del Grupo de Trabajo de Identificadores de Legislación¹⁴, la legislación española se configura a diferentes niveles (estatal, autonómico y local), que se deberían estar relacionar entre sí. Además, cada administración tiene sus propias soluciones técnicas y funcionales, en el sentido de que no sólo estamos hablando de diferentes formatos y estructuras regulatorias, sino también de sistemas de información muy diferentes.

La configuración jurídica española se refleja en el hecho de que existe: una publicación nacional representada por el Boletín Oficial del Estado (BOE); publicaciones regionales representadas por cada uno de los boletines oficiales de cada Comunidad Autónoma; y publicaciones locales donde se encuentran los boletines oficiales provinciales y, además, se podría hablar incluso de la normativa específica de cada ayuntamiento. Si analizamos el estado de los boletines provinciales, la situación es aún mucho más diversa. Esto significa que existe una diversidad entre los diferentes boletines, que afecta a los idiomas de publicación, a la doble publicación, a los sistemas de información, a los servicios y a los metadatos.

En este sentido, cada nivel tiene su propio sistema de publicación de legislación y normas. El BOE se ocupa no sólo de las publicaciones nacionales, sino también de la normativa internacional y, en los casos previstos en los respectivos Estatutos de Autonomía, también de la normativa de un nivel regional superior. Cada Comunidad Autónoma, junto con las ciudades de Ceuta y Melilla, tienen sus propios boletines oficiales. En el caso de la legislación local, se publican en los correspondientes boletines provinciales¹⁵. En total, se tiene un boletín oficial del Estado, 19 boletines de las comunidades autónomas y 43 provinciales.

Por otra parte, esta diversidad resulta estar determinada por otro tipo de factores, como la pluralidad lingüística y la doble publicación. Así, en las comunidades autónomas que tienen varias lenguas oficiales, las diferentes normativas se publican en cada una de sus lenguas. Por otro lado, en el caso concreto del BOE, que también cuenta con 4 suplementos oficiales, algunas disposiciones se publican en diferentes lenguas del Estado. En cuanto a la doble publicación, existen reglamentos que se publican tanto en el boletín regional como en el

14 ELIdata.es-ELI. Identificador Europeo de Legislación. URL: <https://www.elidata.es/>. Consultado el 20 de marzo de 2020.

15 Con excepción de las comunidades uniprovinciales, como Madrid, Murcia y La Rioja, que tienen sus propios boletines.

estatal. Así, las leyes autonómicas también se publican en el BOE, aunque entran en vigor con la publicación autonómica¹⁶.

Todo ello hace que la localización de una normativa concreta se distribuya en muchos casos a través de diferentes sistemas, cada uno con diferentes formas de acceso y estructuración de la información. Por ejemplo, pensemos en el caso de las regulaciones urbanas, donde hay aspectos contemplados en todos los niveles de la administración. Un profesional del sector o una empresa en busca de mejores oportunidades de negocio debería consultar toda esta información distribuida navegando por cada uno de estos sistemas sin que, en muchos casos, esta regulación esté vinculada entre sí.

Todos estos boletines han pasado de la publicación oficial en papel a la publicación electrónica, y la mayoría ofrecen sus contenidos en PDF y HTML, aunque no existe una total uniformidad entre ellos. Así pues, algunos boletines tienen otros formatos como ePub y XML. Además, cada boletín ha decidido su formato de publicación definitivo, por lo que nos encontramos con que algunas publican el boletín en un solo documento que contiene todas las disposiciones, otras la publican en versiones separadas y otras publican ambos tipos de contenido.

Muchos de los sistemas de información se apoyan en bases de datos y plataformas de servicios de contenido que, además, ofrecen servicios de valor añadido a lo que se publica en su boletín oficial (por ejemplo, las bases de datos de la legislación consolidada), difundiendo también información o dando acceso a otras reglamentaciones publicadas en otros boletines oficiales (por ejemplo, una reglamentación regional puede hacer referencia a una reglamentación estatal). Entre los servicios de valor añadido destacan las bases de datos de la legislación consolidada. Sin embargo, en el caso de los boletines oficiales provinciales, se dispone de muy poca legislación consolidada.

En la legislación consolidada, todas las correcciones y modificaciones que han tenido lugar se incorporan a las reglamentaciones originales que, aunque no se consideran oficiales, aportan un valor añadido que facilita la seguridad jurídica. Por ejemplo, un artículo de una regulación sectorial puede ser modificado por otra regulación con poca relación, lo que puede significar que un interesado no conozca este cambio normativo. En este sentido, el BOE y muchas comunidades autónomas publican los textos consolidados, aunque cada uno de ellos sigue sus propias normas de consolidación. Sin embargo, estos textos también se publican en algunos casos por sistemas o portales distintos de los destinados a los boletines oficiales.

16 También hay casos en los que la legislación estatal es publicada por las Comunidades Autónomas, como el Real Decreto de Transferencias, que normalmente se publica primero en el BOE.

Un servicio cada vez más esencial es el análisis jurídico, para lo cual es fundamental que los textos se fragmenten, llegando a nivel de artículos o párrafos, además que estén asociados con términos y materias, con el fin de poder realizar relaciones y clasificarlos en categorías. La existencia de términos compartidos es un aspecto fundamental para poder conocer la normativa publicada en otros boletines. En España, prácticamente todos los sistemas han tomado como base el diccionario EuroVoc¹⁷ para definir sus términos (Alvite Díez, 2012) dentro de sus buscadores, aunque estos sistemas no los han incorporado dentro de los metadatos que describen las normativas. Aunque existe la experiencia de Cataluña que lo han incrustado explícitamente dentro de sus textos (Orenga-Gaya y Giralt, 2011).

Otro servicio que se ofrece es el de los buscadores de legislación. En este caso, el alcance o el tipo de búsqueda de estos motores de búsqueda difiere entre los distintos sistemas de información. Además, en general, cada Comunidad Autónoma tiene en sus publicaciones, en el caso de que las ofrezcan, términos definidos de forma diferente entre ellas, lo que contribuye a una mayor complejidad y dispersión del sistema nacional de publicaciones jurídicas.

En definitiva, dentro del contexto español, hay una serie de servicios destinados a cumplir con las normas de transparencia y reutilización de la información. Estos servicios pueden resumirse en la incorporación de textos consolidados, la trazabilidad, el análisis jurídico y buscadores. A estos servicios los denominaremos **servicios básicos**.

4. Marco teórico

ELI por sí solo no materializa los beneficios de la Web Semántica. Permite que todos los recursos legales tengan un identificador estandarizado, que los datos y sus relaciones se describan con los mismos metadatos y que se hable un idioma común. Sin embargo, esto no es suficiente para que exista una Web Semántica para el ámbito legal. En este sentido, se requiere que se desarrollen servicios sobre los pilares de ELI. Como se ha visto, Francart *et al.* han analizado los servicios ofrecidos por el ELI y sus potencialidades, a estos servicios los hemos clasificado como servicios en red, servicios locales y servicios de provisión, que por otra parte, pueden mejorarse para aprovechar las potencialidades de la características que aporta ELI, no solo para aportar más transparencia sino también para facilitar y potenciar la reutilización.

Por otro lado, dentro del contexto español, hay una serie de servicios, que los hemos denominado servicios básicos, con el objetivo de avanzar en el cumplimiento de los principios de transparencia, como son los servicios de textos consolidados, trazabilidad, análisis jurídico y buscadores.

17 EuroVoc-EUR-Lec. URL: <https://eur-lex.europa.eu/browse/eurovoc.html>. Consultado el 15 de abril de 2020.

Estos servicios tienen que ser implementados y ofrecidos por cada sistema de información, es decir, a nivel local. Pero, de hecho, los servicios finales de la Web Semántica se establecen en red, donde son las interrelaciones entre los diversos servicios ofrecidos en cada sistema de información lo que realmente materializa la Web Semántica. Dicho en otras palabras, los pilares de ELI primero apoyan los servicios a nivel local, pero con el objetivo final de interrelacionarse con otros sitios para trabajar en red. Sobre esta base podemos diferenciar tres arquitecturas diferentes, pero estrechamente relacionadas. Así, tenemos: una arquitectura orientada a servicios, que se apoya en una arquitectura semántica y en una arquitectura tecnológica. Veamos estas tres arquitecturas que serían nuestros modelos de contexto, que nos permitirán definir el modelo completo, es decir, el Marco Teórico.

4.1. Arquitectura de servicios

ELI proporciona una infraestructura tecnológica mínima, lo que facilita que cada sistema de información pueda ofrecer servicios a nivel local y que pueda ofrecer otros servicios en red a otros sistemas de información. Los servicios ofrecidos los hemos clasificado por niveles. El propio ELI ofrece una serie de servicios, soportados por los pilares 1, 2 y 3, que los denominaremos **servicios de infraestructura**. Estos servicios son los que facilitan la identificación, la descripción de los datos y sus relaciones, y un lenguaje común de intercambio. Estos servicios se apoyan en el conjunto de tecnologías descritas en ELI. Además, entre los servicios que deben estar disponibles se encuentran los relacionados con la gestión de las URIs ELI, de manera que se pueda acceder a ellos a través de un sistema guiado a los usuarios.

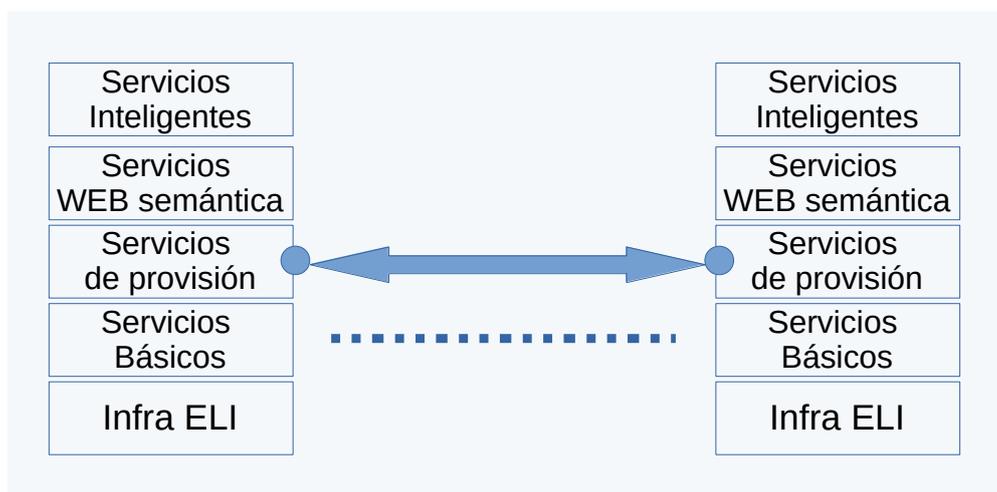
Por otra parte, las unidades de publicación no deben perder de vista que los servicios que ofrecen son servicios de publicación legislativa. Para que la sociedad disponga de la mayor información jurídica posible se deben ofrecer no sólo los textos originales de la legislación, sino también los Proyectos de Ley (así como, otros documentos, como dictámenes, informes, etc.) y los textos consolidados, junto con un análisis jurídico, y también debe haber trazabilidad entre ellos. La principal tecnología que soporta estos servicios son los lenguajes de marcado de las publicaciones, esencialmente HTML y XML. Estos textos también deben relacionarse con los términos apropiados para una mejor clasificación y búsqueda. Hemos denominado a estos servicios como **servicios básicos**.

Las administraciones deben facilitar a la sociedad y a los demás operadores el acceso, la reutilización y la interoperabilidad de los textos legislativos. Para ello, se debe ofrecer un conjunto de **servicios de provisión**. Estos servicios se apoyan en tecnologías como SPARQL y REST y la publicación de conjuntos de datos o *datasets*. Los *datasets* son hojas en las que se describe la información publicada de los datos que se ofrecen. Estos *datasets* se pueden ver como complemento informativo donde se describe y clasifica la información disponible, de esta forma se facilita a los programadores hacer uso de estos servicios.

En realidad, son los servicios en red los que terminarán aportando más valor a la sociedad, especialmente cuando una parte desea conocer información jurídica que trasciende más de un territorio. Estos son los verdaderos **servicios de la Web Semántica**, cuando se establecen relaciones y los datos se entrelazan y se crea una semántica entre ellos. Estos servicios se apoyan en redes y vocabularios controlados y compartidos de conocimiento. Un sitio puede ofrecer estos servicios en red consumiendo servicios de otros sitios, o pueden comportarse como un simple proveedor de servicios.

Finalmente, con los **servicios inteligentes**, aplicando tecnologías de inteligencia artificial y de procesamiento de lenguaje natural, que permiten adquirir datos y metadatos, no descritos explícitamente, con el fin de enriquecerlos y establecer nuevas relaciones, es cuando la Web Semántica será realmente más completa y se aproxime al razonamiento humano.

Figura 1. Esquema del Marco Teórico



Fuente: Elaboración propia.

Estos servicios siguen una arquitectura de red donde se concretan los servicios locales y de red para obtener la arquitectura de servicios que se representa en la figura 1. Tal como se representa en esta figura, los servicios de provisión son los puntos de conexión entre los diferentes sistemas. Por encima de los servicios de aprovisionamiento se pueden desarrollar servicios de red, en el mismo sitio o en otros sitios.

4.2. Arquitectura semántica

Hemos visto que los metadatos ELI no son suficientes para lograr una completa WEB semántica legal. No son suficientes porque no incluyen todos los términos y relaciones que describen un texto basado en aspectos como los términos específicos de un sector se refiere o si se trata de derechos fundamentales. Los metadatos ELI se centran más en la descripción de los conceptos y relaciones asociados a los documentos, como los relacionados con el tipo de texto, es decir, si se trata de una Ley, un Reglamento, etc., o las relaciones entre ellos, como las relaciones de consolidación y modificación.

Cuando un texto se asocia con términos más ricos es cuando realmente se pueden establecer relaciones más ricas. Para ello es necesario complementar la ontología ELI con vocabularios más ricos en conceptos legales. Solo si estos vocabularios son controlados y compartidos es posible establecer conversaciones entre sistemas de información. Para facilitar el uso de los vocabularios se han desarrollado diversas tecnologías específicas definidas por el W3C. Como ejemplos de estas tecnologías, tenemos el *Data Catalog Vocabulary* (DCAT), el *Vocabulary of Interlinked Datasets* (VOID) y el Simple Knowledge Organization System (SKOS). Además, los vocabularios específicamente jurídicos se pueden complementar con otros vocabularios RDF, como el *Friend Of A Friend* (FOAF), que permite, entre otras, definir relaciones de muy diverso tipo entre personas. Dado que estos vocabularios y tecnologías se describen con RDF, pueden evolucionar con la ontología ELI.

Cada sistema de información suele partir de sus propios vocabularios, donde, en general, parten de sus propios términos que describen aspectos jurídicos, pero que son específicos de cada lugar. De esta manera tienen un vocabulario controlado pero no compartido. Este vocabulario puede incluirse en las tecnologías anteriores para ofrecer servicios locales, integrándose con los motores de búsqueda que utilizan SPARQL. En este caso, se dispondría de un vocabulario controlado RDF, que permitiría ofrecer redes de conocimiento jurídico más ricas a nivel local, pero no permitiría la interoperabilidad completa con otros sistemas, ya que no comparten la misma semántica.

Por lo tanto, se requiere una evolución para pasar a un vocabulario controlado y compartido. Esto puede hacerse, tal como describen Francart *et al.*, utilizando un vocabulario como el EuroVoc o utilizándolo como un pivote. En cualquier caso, parece que EuroVoc se está convirtiendo en el vocabulario de referencia estándar para complementar la ontología ELI. En este sentido, es necesaria una estrategia de evolución de los vocabularios actuales hacia este vocabulario compartido. Una estrategia a seguir podría ser la de definir un vocabulario controlado y compartido a nivel de español. Este podría ser un pivote entre los sistemas de información en

España, y luego este último podría utilizar el vocabulario EuroVoc como un pivote para permitir una interoperabilidad semántica más rica a nivel europeo.

4.3. Arquitectura tecnológica

En los dos apartados anteriores se han ido nombrado las distintas tecnologías relacionadas con los servicios y la semántica, a continuación estas tecnologías se ordenarán en una arquitectura tecnológica. Se ha visto que las tecnologías descritas en la especificación ELI no son las únicas, aunque son las más importantes y podemos considerarlas las tecnologías clave para tener datos abiertos entrelazados y servicios de web semántica. Gran parte del potencial de estas últimas tecnologías es que almacenan y describen los datos dentro de sus propios archivos. Al mismo tiempo, se establecen relaciones entre todos los datos que se distribuyen a lo largo de la WEB.

Los tres pilares ELI proporcionan la base sobre la que construir una web semántica legal. Esto permite navegar a través de los archivos para acceder y representar el conocimiento distribuido. Entre los conceptos clave es que manejan metadatos comunes. Esto es proporcionado por el Pilar 2. Pero, además, se necesita un lenguaje común para permitir la interoperabilidad, que es proporcionada por el Pilar 3 a través de RDFa (y otros formatos) que estructuran los datos. ELI inicialmente estableció RDFa y ahora también incluye JSON-LD, que facilita la implementación, ya que requiere menos esfuerzo por parte de los desarrolladores. Por otra parte, aunque no está incluido en la especificación del ELI, TURTLE se está consolidando como un lenguaje para el modelado de información, ya que se considera más fácil de leer y editar manualmente que los formatos basados en XML. Pero también necesita que OWL describa las relaciones semánticas más ricas, y por lo tanto, que navegue de manera más inteligente entre los diferentes archivos distribuidos. SPARQL es necesario para poder consultar esta base de datos distribuida utilizando un lenguaje que se adapte a esta estructura de archivos.

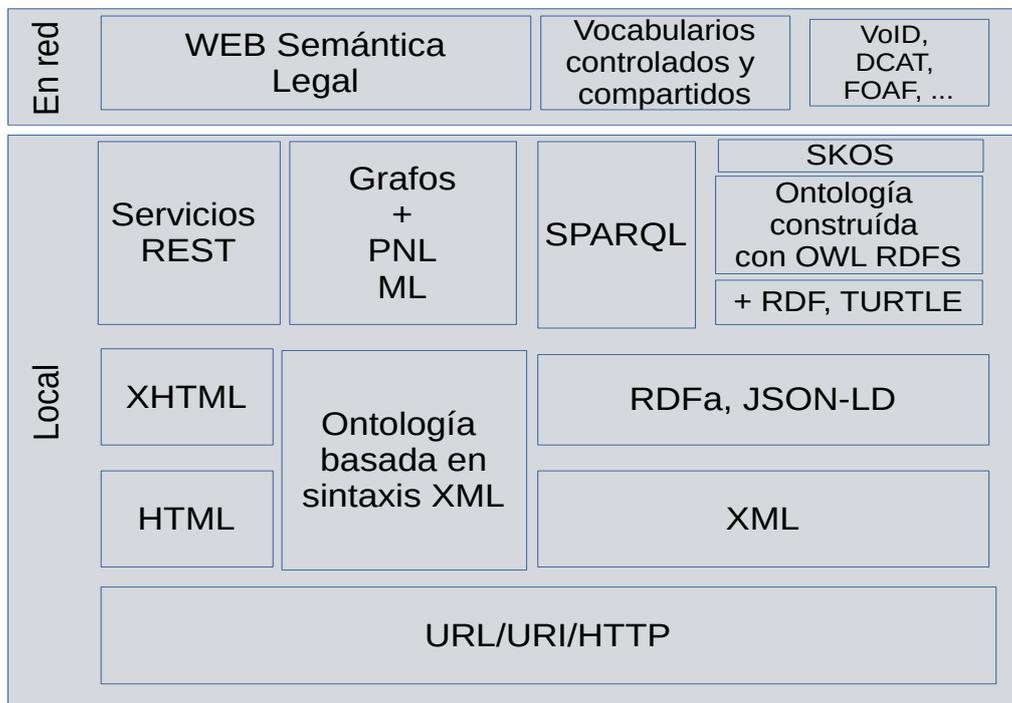
Con la combinación de estos componentes, tenemos la capacidad de codificar la semántica y crear un razonamiento automatizado, vinculando e intercambiando información entre las diferentes fuentes y su gestión. Con esta arquitectura, las ontologías que describen el mundo de las publicaciones legislativas, todas las entidades juegan un papel vital para hacer realidad la WEB Semántica. Pero, si se quiere enriquecer este conocimiento es necesario incluir más vocabularios disponibles en la WEB Semántica junto con vocabularios legales controlados y compartidos, como el EuroVoc y otros recomendados por el W3C, descritos anteriormente en la arquitectura semántica. En este sentido, los sitios locales deberían utilizar el SKOS que facilita la conexión en red con los diferentes vocabularios.

Por otra parte, dentro de un entorno local, para facilitar el uso por parte de los programadores de otros entornos externos, los servicios REST son los más extendidos. De esta manera, con un servicio REST en el sitio local, utilizando, por ejemplo, JSON, se puede acceder a los contenidos en XML a través de HTTP. Además, para facilitar la interoperabilidad, el sitio local debe complementarse con *datasets* y validadores que proporcionen más información y calidad para los reutilizadores o los servicios de la Web Semántica. Esto proporcionaría una completa arquitectura tecnológica que cubre las muchas potencialidades del ELI.

Para trabajar en red, entre los servicios que se pueden proporcionar tenemos la generación de grafos de conocimiento legal, que se generan en local, pero usan datos externos. Estos pueden ser ofrecidos a nivel local, sobre los recursos que están disponibles localmente, incluso enlazando con otras relaciones que se establecen con otros sitios. Finalmente, cuanto existen vocabularios legales y compartidos en red, y de manera generalizada en otros sitios, se pueden generar redes de conocimiento jurídico que involucren a todos los sitios. Todo ello puede enriquecerse con servicios inteligentes que utilicen principalmente el Procesamiento del Lenguaje Natural (PNL) y *Machine Learning* (ML) no solo para obtener información no explícita, sino para realizar consultas en lenguaje humano que se traduzca a preguntas entendibles por máquinas.

En la figura 2 se representa la pila de todas estas tecnologías, y como estas se relacionan entre sí. Por último, indicar que esta arquitectura tecnológica no se considera que sea un mínimo, dado que, tal como se establece en la especificación ELI, cada Estado podrá escoger aquellos aspectos que mejor se adapten a las peculiaridades ya sean técnicas o de su configuración jurídica. Tampoco, se puede considerar que sea un máximo: la riqueza de tecnologías relacionadas con LOD es bastante amplia con lo que cada sistema pueda incorporar otras tecnologías o innovaciones que faciliten sacar sus mejores beneficios para mejorar la información jurídica.

Figura 2. Arquitectura tecnológica.



Fuente: Elaboración propia.

4.4. Modelo completo

Las tres arquitecturas anteriores están estrechamente relacionadas entre sí. Así, tenemos: una arquitectura orientada a servicios, que se apoya tanto en una arquitectura semántica como en una arquitectura tecnológica. Cuando integramos las tres arquitecturas anteriores tendremos un modelo completo, que conducirá a nuestro Marco Teórico. Esta integración significa la interposición de las tecnologías con los servicios considerando los aspectos semánticos. En la tabla 1 proponemos un modelo por capas de los servicios, donde los niveles 1, 2 y 3 se corresponden con los servicios de infraestructura.

Tabla 1. Modelo Completo

Nivel	Denominación	Descripción
1	Pilar 1	Se identifican los recursos a través de las URIs y se gestiona las URIs.
2	Pilar 2	Se definen los metadatos.
3	Pilar 3	Se definen las tecnologías o formatos LOD.
4	Servicios básicos	Se definen los servicios de publicación, análisis jurídico y de trazabilidad de los textos.
5	Servicios de provisión	Se definen los servicios para la accesibilidad, reutilización e interoperabilidad.
6	Servicios en red	Se definen servicios que permiten en establecimiento de relaciones en red, como son los grafos de conocimiento.
7	Servicios inteligentes	Se definen tecnologías que aportan más valor semántico a partir de relaciones implícitas en los textos.

Fuente: Elaboración propia.

La implementación de todos los servicios y tecnologías anteriores no puede considerarse como una solución única desde el principio, es decir, la solución final se debe implementar de forma gradual. Por otra parte, cada una de las administraciones parte de sistemas diferentes y de prioridades distintas, lo que significa que la implementación entre ellos no tiene que ser con las mismas hojas de ruta. Por lo tanto, es interesante tener un modelo que permita catalogar el estado de cada uno. En este sentido, hemos definido etapas o estados de madurez específicos para este trabajo. Una etapa o estado de madurez es una etapa definida, representada por un escalón, donde se recoge la etapa evolutiva en la que se encuentra cada sistema de información. Así, cada etapa de madurez ofrece una base para la mejora continua de la implementación de servicios para una mejor información jurídica.

En el caso de España, podemos partir de una etapa inicial que estaría asociada solo con la especificación ELI española (Fase 1). En esta fase, el Pilar 3 no es obligatorio, por lo que la interoperabilidad no estaría garantizada. Además, la descripción de los metadatos sólo incluiría los considerados obligatorios, que no son suficientes para proporcionar una descripción completa de los datos y sus relaciones. Este estado sería el punto de partida para evolucionar a estados superiores. Por lo tanto, se propone un **estado básico**, que incluiría el Pilar 3. En esta etapa se dispondría de los textos originales, apoyados en los tres pilares, con un nivel de descripción de los metadatos superior al de los únicos metadatos obligatorios. Este estado

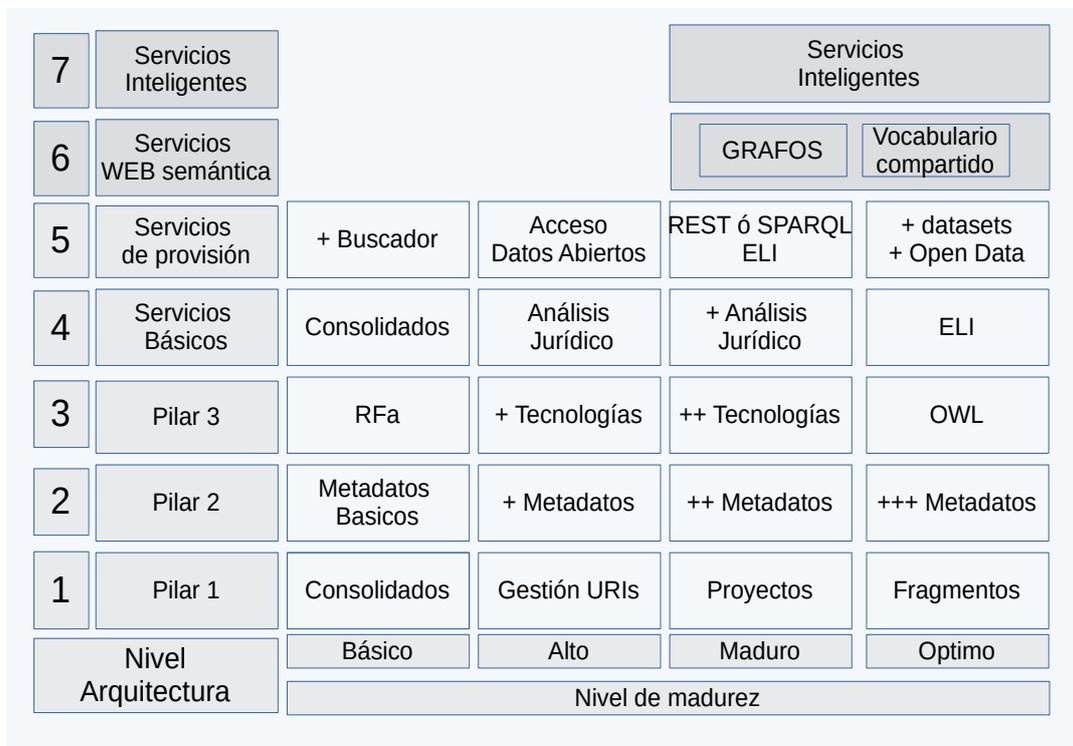
garantizaría la interoperabilidad, aunque no sería rica, ya que no sería posible seguir la fase en que se encuentra una norma. En este caso, los metadatos comenzarían a incluirse dentro de los metadatos no obligatorios de la especificación ELI española. Además, debe haber servicios de motores de búsqueda (no necesariamente apoyados por ELI) que permitan búsquedas por términos o temas.

En la siguiente etapa, el **estado alto**, se incluirían más metadatos, incluidos los relacionados con los textos consolidados, y los metadatos ELI se relacionarían con la cronología de las regulaciones, conociendo el vigor de cada regulación. Para lo cual sería necesario incluir algunos metadatos de la especificación europea ELI no incorporados en la especificación española. Además, el Pilar 3 evolucionaría para tener otros formatos que facilitarían la interoperabilidad, como los formatos JSON-LD y TURTLE. El análisis jurídico sería más completo, en el que comenzarían a incorporarse servicios como el análisis de la vigencia, la trazabilidad o las relaciones con materias o términos.

En una etapa superior, el **estado maduro**, estos últimos servicios serían más completos. Además, en este estado se deberían ofrecer datos abiertos, no necesariamente apoyados por ELI. De esta manera, sería posible rastrear los Proyectos (o dictámenes u otra documentación) hasta los textos consolidados, incluidos los textos originales, para los que se deberían incorporar más metadatos. Además, estarían incorporados de forma completa, servicios REST o SPARQL para la accesibilidad, la reutilización y la interoperabilidad basados en ELI. Esto facilitaría a otros operadores el acceso a la información, así como la interoperación con otros sistemas jurídicos de manera más sencilla.

El **estado óptimo** incluiría más identificadores ELI, para los cuales se tienen que incorporar más metadatos para establecer relaciones a nivel de fragmentos. Además, los metadatos deberían permitir relaciones más ricas que se vinculen a vocabularios controlados y compartidos. Se podría incluir un OWL propio para establecer relaciones más ricas entre los textos. En esta etapa, el análisis jurídico se apoyaría en OWL. Tanto los servicios REST como los SPARQL ya estarían incorporados con suficientes *datasets* y sistemas de validación para permitir una interoperabilidad de calidad.

Figura 3. Esquema del Marco Teórico



Fuente: Elaboración propia.

Una vez que se alcancen el estado maduro y óptimo, permitiría una evolución para enriquecer los servicios de cada sistema. Se incorporarían términos y vocabularios, con relaciones entre metadatos compartidos, y que permitirían la interoperabilidad utilizando tecnologías como VOID y DCAT nombradas en la arquitectura semántica, que facilitan el acceso en red a los distintos vocabularios de los diferentes sitios. De esta manera, las redes de conocimientos jurídicos pueden incorporarse a nivel local y ampliarse a nivel nacional en una red. Si la ontología ELI no incorpora un diccionario compartido, entonces tendría que desarrollarse a nivel nacional. A este nivel, se pondría poner en marcha la inclusión de servicios de red inteligente.

Con este modelo completo se habría definido el Marco Teórico. Este Marco Teórico descrito se esquematiza en la figura 3.

5. Metodología

La metodología para describir el estado actual en España se ha basado en la observación de las distintas implementaciones de soluciones en los portales de los distintos sistemas de información. En nuestro caso, se ha aplicado: la selección de las muestras, la recopilación, la reducción, la representación, y la interpretación de información.

Se inicia seleccionando las muestras a partir del universo de estudio. Estas serán las implementaciones de las distintas administraciones que disponen de una solución implementada de la especificación ELI española. Por tanto, solo se han analizado los sistemas del Estado, Cataluña, Murcia, País Vasco y Valencia, que son los que tenían implementado ELI en marzo de 2020. El resto o están en fase de implantarlo o no lo han iniciado a pesar de la obligación que tenerlo implantado antes de 2020.

La recopilación se trata como trabajo de campo, donde se plantea una observación directa sobre las soluciones de todos los portales WEB de la muestra seleccionada. La guía para definir la recogida se hecho en base al Marco Teórico generado. La observación directa se ha basado en la navegación en cada portal y el análisis de los códigos en HTML de textos, donde se incrustan los metadatos y la ontología ELI. En la figura 4 se puede ver un ejemplo de parte de un código HTML con metadatos ELI, *donde se puede observar tres metadatos: title, publisher y date_publication.*

Figura 4. Extracto de código HTML con metadatos ELI.

```
<meta about='https://www.boe.es/eli/es-cb/l/2019/03/07/2/dof/spa' property='http://data.europa.eu/eli/ontology#title' content='Ley 2/2019, de 7 de marzo, para la igualdad efectiva entre mujeres y hombres.' datatype='http://www.w3.org/2001/XMLSchema#string' />
<meta about='https://www.boe.es/eli/es-cb/l/2019/03/07/2/dof/spa' property='http://data.europa.eu/eli/ontology#publisher' resource='www.boe.es' />
<meta about='https://www.boe.es/eli/es-cb/l/2019/03/07/2/dof/spa' property='http://data.europa.eu/eli/ontology#date_publication' content='2019-03-29' datatype='http://www.w3.org/2001/XMLSchema#date' />
```

Fuente : <https://www.boe.es/eli/es-cb/l/2019/03/07/2/con>.

Consultado el 18 de mayo de 2020.

La reducción se ha hecho a través de un análisis comparativo de las distintas soluciones recopiladas, delimitando la información que se obtiene en la fase de recogida de datos. Se hace para que sea más manejable, para ello se va categorizando cada sistema en función de los distintos aspectos que recoge el Marco Teórico. Así, se ha sistematizado la información en elementos diferenciados y representándolos en base a resúmenes comparativos de todos los sistemas analizados. La representación final resume las distintas soluciones para cada aspecto que se recoge en diferentes plantillas basadas en el Marco Teórico. Esta

representación final de los resultados se hace en forma de tabla de doble entrada: el nivel del Marco Teórico y para cada uno de los sistemas analizados.

La interpretación de la información se asocia con la discusión de los resultados. Para valorar cada sistema se ha asociado una puntuación del 1 al 5 en función del estado de madurez, pero asociando un estado de madurez por cada uno de los niveles del Marco Teórico. Así, se realiza un análisis detallado de cada nivel, y la puntuación final de cada sistema es la media de la puntuación para los distintos aspectos de cada nivel. En esta puntuación final, no se ha tenido en cuenta los niveles 6 y 7, dado que se tratan de niveles que trabajan en red y, aparte de que no existen todavía servicios en red, el objetivo es determinar el estado de cada sistema en sus entornos locales. En definitiva, se han mapeado y puntuado los aspectos definidos en el Marco Teórico con los resultados obtenidos.

6. Resultados y discusión

Como se ha dicho, la recogida de datos y su análisis se ha definido en base al Marco Teórico; y solo se han analizado aquellos sistemas que tienen implantado ELI a fecha de marzo de 2020, cumpliendo con especificación ELI española (Fase 1) (Estado, Cataluña, el País Vasco, la Región de Murcia y Valencia). Además, se ha tomado la especificación ELI española como referencia para las etapas iniciales. En esta especificación se ha definido unos requerimientos mínimos iniciales para garantizar que se disponga de una base técnica que permita la interoperabilidad a corto plazo entre estos sistemas. También, se ha visto que estos sistemas han ido algo más allá de esta especificación incorporando el Pilar 3 y algunos sistemas han incluido metadatos de la especificación europea que no estaban reflejados en la especificación española. En cualquier caso, el alcance de cada implantación ha sido diferente.

Por otra parte, cada una de las Comunidades Autónomas analizadas tienen configuraciones diferentes para ofrecer la información jurídica. Como es normal, todas disponen del servicio de boletín oficial donde se publica las normas originales (oficiales y auténticas), pero difieren en la forma de ofrecer servicios con más información jurídica. El Estado y Valencia en el portal del boletín se ofrecen todos los servicios asociados con los datos abiertos y ELI, integrados en el mismo portal, aunque el Estado desarrolla de forma más exhaustiva los servicios de análisis jurídico, suplementándolos con información jurídica de códigos, dictámenes, jurisprudencia, etc. Cataluña y el País Vasco¹⁸ han creado portales específicos para ofrecer estos servicios de análisis jurídico y ELI, donde Cataluña ofrece la trazabilidad de los proyectos y dictámenes previos a las normas asociadas a las publicaciones. Murcia, a excepción de los textos consolidados que se ofrecen a través del Portal de Transparencia de su Comunidad Autónoma, se ofrecen estos servicios en el mismo portal del boletín.

18 Cataluña dispone de Portal Jurídico (<https://portaljuridic.gencat.cat/>) y el País Vasco del Portal Legegunea (<https://www.legegunea.euskadi.eus/>).

Relacionado con el Pilar 1, los servicios de gestión de las URIs, los cuales facilitan el acceso a las publicaciones en base a identificadores ELI en el mismo portal, se han implementado en base a dos formas distintas: a través de carpetas, donde se accede a estas navegando a través de los mismos identificadores ELI de la URI; y a través de servicios basados en buscador. No es directo conocer cuál de ambas formas es mejor de cara a la sociedad, aunque, en cualquier caso, se puede considerar que son complementarios. Cada sistema ha optado en aplicar uno de los dos, pero no ambos. Posiblemente, para una mejor accesibilidad se puede plantear que los distintos sistemas evolucionen a una solución que contemplen ambas soluciones. Por último, no todos los sistemas disponen de los mismos identificadores. Así, nos encontramos que algunos incluyen la normativa consolidada dentro su URI ELI, como el identificador */con/* (Estado, Cataluña, Valencia), o Cataluña que incluye los identificadores para los sumarios (*/sum/*) y para el boletín completo (*/dia/*).

En los metadatos del Pilar 2, en general, se han ampliado los metadatos de la especificación española. Entre estos metadatos incorporados podemos diferenciar diversos tipos: de información, *id_local*, *legal_value*, *title_short*, *title_alternative*, *licence*, *rightsholder_agent* (Estado, Cataluña, País Vasco); de trazabilidad, *responsibility_of*, *published_in_format* (Cataluña, País Vasco); de relaciones con materias, *is_about*, (Cataluña, País Vasco); de relaciones con otros textos, *changes*, *repeals* (País Vasco); y de vigor *in_force*, *date_no_longer_in_force*, *date_applicability* (Cataluña, País Vasco).

Con respecto al Pilar 3, el Estado y Cataluña ha incluido RDF en XML. Cataluña y País Vasco han ampliado XHTML para ofrecer otros formatos LOD, pero lo han hecho de forma diferente. Así, Cataluña ofrece directamente ficheros de intercambio junto con cada texto, para ello, se pueden descargar los ficheros RDF y TURTLE en la publicación de cada texto en el Portal Jurídico, y se ofrecen junto a los formatos PDF y XML. Este último sigue el estándar *Akoma Ntoso*¹⁹. País Vasco ofrece los ficheros de intercambio en su portal Open Data²⁰, y se ofrecen junto con *datasets* usando RDF en TURTLE.

En cuanto a los servicios básicos, los distintos sistemas disponen de servicios de análisis jurídico y de búsqueda. La trazabilidad principalmente abarca a los textos consolidados relacionados con los textos originales, aunque con distintos niveles de desarrollo. Sobretodo Cataluña permite una auténtica trazabilidad de los Proyectos legislativos hacia el texto original, incorporando además en la trazabilidad a los dictámenes junto con la documentación asociada²¹. Este tipo de

19 Akoma Ntoso fue promovido inicialmente por las Naciones Unidas para publicar las leyes de los parlamentos africanos, junto con los documentos legislativos de apoyo (Barabucci *et al.*, 2009).

20 Open Data Euskadi. URL: <http://opendata.euskadi.eus/>

21 Un ejemplo de trazabilidad, se puede ver en en cuadro relación cronológica del Portal Juridic de Catalunya: URL: <https://portaljuridic.gencat.cat/eli/es-ct/L/2010/08/03/26> . Consultado el 21 de mayo de 2020.

servicios tiene sus peculiaridades y se relacionan estrechamente con servicios de edición. De forma general no se hace uso de los metadatos ELI para realizar la trazabilidad, sino que se hace a nivel de aplicación.

En cuanto a los servicios de provisión, estos están directamente relacionados con las tecnologías REST y SPARQL y los datos abiertos. En el primer caso, los cinco sistemas, en general, o no se han terminado de desarrollar o no aportan ninguna innovación nueva con respecto a experiencias europeas. En el caso de los datos abiertos, sobre todo el País Vasco los desarrolla en base a ELI en su portal Open Data y, además, se comunican a nivel nacional y europeo, mejorando la información jurídica. Por último, indicar que no se puede considerar que todavía existan servicios de WEB semántica y servicios inteligentes en ninguno de los sistemas.

De los resultados obtenidos durante el periodo de marzo a mayo de 2020 a través de la recogida de datos y su análisis tenemos que la situación de cada sistema se puede resumir en la tabla 2.

Tabla 2. Valoración del estado actual de cada sistema

Sistema	Nivel 1	Nivel 2	Nivel 3	Nivel 4	Nivel 5	Valoración
Estado	2	1	2	2	2	Alto (1,8)
Cataluña	3	4	3	3	2	Maduro (3)
Murcia	1	1	1	2	2	Básico (1,4)
País Vasco	1	4	3	3	1	Maduro (3)
Valencia	2	1	1	2	1	Básico (1,4)
Valoración	Alto (1,8)	Alto (2,2)	Alto (2)	Alto (2,4)	Alto (2,2)	Alto (2,1)

Fuente: Elaboración propia.

A continuación se procede a la discusión de los resultados, diferenciando el análisis por cada sistema y por cada uno de los niveles.

6.1. Discusión por cada sistema

Como se ve en la tabla 2, la calificación global de los cinco sistemas tenemos una valoración de Alto. Esto quiere decir que la implantación de ELI de los sistemas que lo han acometido han ido más allá de los requisitos obligatorios de la especificación española. Además, los servicios que actualmente están ofreciendo a la sociedad son lo suficiente maduros para a partir de este momento enfocarse a la implantación en servicios con más valor basados en los conceptos de datos abiertos entrelazados. En cualquier caso, en dicha tabla podemos ver que la situación actual es diversa: tenemos dos sistemas (Cataluña y el País Vasco) con un estado de madurez calificado de Maduro; uno de Alto (Estado) y dos Básicos (Murcia y Valencia). Por otra parte, aunque tenemos estos dos sistemas clasificados como Maduro, ambos sistemas difieren entre sí.

Cataluña se ha enfocado en ofrecer mejores servicios relacionados con los servicios básicos. De esta forma, de los dos, es el sistema que ofrece la mejor solución que aporta un mejor análisis jurídico junto a la trazabilidad, donde además se incluye el etiquetado de la vigencia a través de ELI y una trazabilidad completa, que no solo se queda en los Proyectos, sino que va más allá, incluyendo los dictámenes y su documentación asociada. Además, ha iniciado la incorporación de los Proyectos legislativos a una filosofía de identificación con ELI, lo que permite a este sistema ofrecer en el futuro servicios de provisión para adquirir información con mayor trazabilidad.

El País Vasco ha optado por una estrategia diferente, y su estado como Maduro se debe a que ha acometido de forma completa los servicios de provisión. Por otra parte, aunque tiene cierta trazabilidad desde los Proyectos legislativos y se etiqueta los estados de vigencia de las normas, su solución no es tan completa como la de Cataluña. En el caso del País Vasco, los servicios de provisión no son acometidos directamente por las unidades de publicación de su boletín oficial y de su portal de información jurídica (Portal Legegunea), sino que se ofrecen a través del portal de datos abiertos del Gobierno Vasco. Esto probablemente es la causa del éxito de disponer de servicios de provisión completos, ya que el esfuerzo que supone implantar este tipo de servicios ha sido canalizado a través unidades especializadas.

Los otros tres sistemas todavía no han alcanzado estados de madurez suficientes que supongan dar un salto a corto plazo para ofrecer servicios de datos abiertos entrelazados. El Estado está clasificado como Alto gracias a su madurez en servicios básicos de análisis jurídico y que ofrecen servicios de provisión pero sin estar estos basados en ELI, sino en servicios REST asociados a ficheros XML²². Pero al igual que los otros dos sistemas, clasificados como Básico, disponen de una solución basada en ELI que han ido más allá de los requisitos obligatorios. Ahora, estos los tres sistemas tendrán que optar por una estrategia en que evolucionen hacia una mejora de los servicios básicos o una estrategia que apuesten por los servicios de provisión.

6.2. Discusión por cada nivel

La clasificación global del estado por cada uno de los niveles parece más homogénea. Aunque esto principalmente también se debe a que es una valoración media donde estados con más altos niveles de madurez se compensan con niveles inferiores.

En el caso de Nivel 1 la clasificación global es Alto. Ninguno ha acometido la incorporación de identificadores asociados con la fragmentación de los textos, lo que supone que no exista ninguno con clasificación Óptimo. Pero sobre todo un

22 El servicio de datos a abiertos del BOE se puede consultar en BOE.es – Datos Abiertos. URL: <https://www.boe.es/datosabiertos/>. Consultado el 17 de abril de 2020.

sistema (Cataluña) ha iniciado el proceso de etiquetar los proyectos basándose en ELI, por lo que solo en este caso se ha clasificado como Maduro. Este es el motivo de disponer de una clasificación global como Alto. Para avanzar en este nivel se requiere que los distintos sistemas empiecen a incorporar identificadores ELI a los Proyectos legislativos y dictámenes²³. La incorporación de etiquetado de fragmentos posiblemente suponga un esfuerzo alto. Aquellos sistemas que ya tienen una fragmentación en sus textos originales o consolidados en sus textos publicados en HTML son los que están en mejor posición para incorporarlos. Aunque en este caso, para evitar que cada sistema defina identificadores distintos sería conveniente que se empiece a definir una solución homogénea entre ellos.

En el caso de Nivel 2 la clasificación global es Alto. Lo que supone que, en general, los sistemas disponen de metadatos ELI con información al menos sobre uno de los siguientes aspectos: vigencia, trazabilidad o relaciones. La realidad es que, sobre todo, dos sistemas (Cataluña y País Vasco) disponen de información de los tres aspectos, por lo que se han valorado como Óptimos. El resto (Estado, Murcia y Valencia) no tienen metadatos sobre estos tres aspectos. En este sentido, estos tres sistemas deberían ir avanzando en incorporar más metadatos de los que se definen en la especificación ELI española.

En el caso de Nivel 3 la clasificación global es Alto. En este caso Murcia y Valencia se han quedado en la publicación de XHTML con RDFa. El resto han avanzado en incorporar más formatos, aunque el Estado se ha quedado en publicar XML/RDF y solo Cataluña y el País Vasco publican más formatos (JSON-LD, TURTLE y RDF)²⁴. El disponer de más formatos es esencial para poder generar en el futuro grafos de conocimiento, por ello los tres sistemas más rezagados deberían incluir generar otros formatos.

En el caso del Nivel 4 la clasificación es Alto. En este caso, ninguno de los cinco sistemas han avanzado en incorporar ELI de forma generalizada en el análisis jurídico. Cataluña y el País Vasco son los que están en mejor posición dado que tienen definidos e implementados metadatos ELI de información sobre vigencia, trazabilidad y relaciones, lo que es fundamental para avanzar en incorporar ELI en el análisis jurídico. En este nivel, Cataluña es el que dispone de servicios WEB que permiten una trazabilidad más completa, aunque todavía no ha establecido relaciones ELI dentro de sus metadatos entre los Proyectos, originales y consolidados. Disponer de metadatos más ricos y las relaciones establecidas a nivel de aplicación no debería llevarle mucho esfuerzo. Eso sí, dado que Cataluña

23 Se puede ver un ejemplo de identificador basado en ELI para dictámenes, del Portal Jurídico de Cataluña, en la dirección [https://portaljuridic.gencat.cat/pjcdic/es-ct/dic/2010/06/09/\(1\)](https://portaljuridic.gencat.cat/pjcdic/es-ct/dic/2010/06/09/(1)). Consultado el 7 de febrero de 2021.

24 El código JSON-LD se incrusta dentro de las publicaciones HTML. Como ejemplos de ficheros TURTLE y RDF, se pueden ver en su portal Open Data <https://opendata.euskadi.eus/catalogo/-/legislacion-de-la-comunidad-autonoma-de-euskadi-bopv-european-legislation-identifier-eli/>. Consultado el 16 de mayo de 2020.

no ha avanzado tanto como el País Vasco en los servicios de provisión, no se beneficiaría de las potencialidades que supone en incorporar ELI en los servicios básicos. El caso del País Vasco, le faltaría mejorar los servicios de trazabilidad para equipararse a la solución de Cataluña para sacar más provecho a los servicios de provisión que tiene implementados. El resto deberían avanzar en dos aspectos: avanzar en la incorporación completa de metadatos ELI de vigencia, trazabilidad y relaciones; y mejorar sus servicios básicos incluyendo la trazabilidad completa, ya que sus análisis jurídicos fallan en este aspecto.

En el caso del Nivel 5 la clasificación es Alta. En este caso, sobresale el País Vasco que tiene implantado servicios de provisión basados en ELI usando para ello su Portal de Datos Abiertos²⁵. La clasificación del resto es Alta, con la excepción de Valencia, ya que dan servicios de provisión, aunque no basados en ELI, por lo que sus esfuerzos deben ir encaminados en incorporar tecnologías LOD a las publicaciones legislativas, ya sea a través de portales específicos de datos abiertos o en sus propios sistemas.

En el caso de los niveles superiores, 6 y 7, ninguno de los sistemas está ofreciendo grafos de conocimiento, por ello no se incluyen en la tabla 2. El País Vasco tiene las tecnologías implementadas para poder ofrecer este tipo de servicios. En este sentido, sería interesante que empiece a abordar pilotos que avancen en la implantación de este tipo de servicios a nivel local. A nivel en red todavía no existen al menos tres sistemas que estén en la actualidad preparados para ofrecerlos. En este caso, se considera que el Estado debería ser el que lidere este tipo de servicios en red para que se vayan incorporando otros sistemas, es decir, se considera que el Estado debería ser el que realice los principales esfuerzos en preparar sus sistemas incorporando servicios de provisión basados en ELI y más metadatos con más información. Por otra parte, existe la problemática de que no existe un diccionario compartido de términos o materias incrustados en la ontología ELI. En general, los sistemas trabajan con EuroVoc, aunque parece que solo Cataluña la incorporado en su ontología. En cualquier caso, un diccionario compartido debe estar al menos definido como mínimo a nivel nacional y que este pueda hacer de pivote con el europeo.

En definitiva, todavía se está lejos que se puedan ofrecer servicios en red entre sistemas nacionales. Sobre todo, dos sistemas, Cataluña y el País Vasco, están en una mejor posición para evolucionar hacia el estado Óptimo, es decir, en ofrecer servicios en red o generar a nivel local servicios de grafos de conocimiento. Por tanto, entre los aspectos clave para poder servicios en red, es que se extienda la implantación de ELI, se creen más servicios sobre ELI, y al mismo tiempo, exista mayor uniformidad entre los sistemas.

25 Como ejemplo de servicios SPARQL se puede acceder a través de su portal Open Data <http://api.euskadi.eus/sparql/>. Consultado el 18 de mayo de 2020.

7. Conclusiones

El tema de este trabajo se encuadra dentro del campo legal y los datos abiertos entrelazados aplicado a las publicaciones oficiales e información jurídica. Una de las principales aportaciones a este respecto se refiere a sistematizar, categorizar y completar los servicios que aporta ELI relacionándolos con un conjunto amplio de tecnologías de la WEB Semántica. De esta forma, se ha obtenido un marco que integra la visión servicios con la visión tecnológica. Otra de las principales aportaciones de este trabajo se refiere al análisis de la implantación de ELI en España. En este sentido, este trabajo ha hecho un recorrido y análisis exhaustivo por las distintas implementaciones de ELI que se han realizado en España.

En la actualidad la implantación de ELI, tanto en Europa como en España, no se puede decir que de forma general se haya mejorado la información jurídica, dado que los beneficios de ELI todavía no se ven reflejados en mejores servicios de información jurídica. Estos se prestan a través de los servicios básicos, los cuales actualmente tiene una baja involucración de ELI, y prestándose, fundamentalmente, sin el apoyo de ELI a través de la navegación y tratamiento manual en sus portales. En cualquier caso, en el estado actual, los pocos sistemas que han implantado ELI han creado las bases para avanzar en que la información jurídica pueda ser mejorada a través del procesamiento automático. Por otra parte, mientras existen sistemas como el de Cataluña, el País Vasco y del Estado que según los resultados de este trabajo parecen tener una línea clara sobre cómo avanzar en la implantación de ELI, existen todavía sistemas que ni siquiera han abordado su implantación.

En cuanto a las aplicaciones de este trabajo, tenemos que puede ir destinado a los responsables de las distintas unidades encargadas de la publicación de los textos legislativos, en el sentido que tienen una arquitectura conceptual tanto de servicios como tecnológica de los aspectos que se tienen que abordar en sus sistemas. Estamos hablando no solo del resto de las Comunidades Autónomas que están en su fase inicial o no lo han iniciado, sino de un total de 43 boletines oficiales provinciales. Se ha podido ver en este trabajo que la implantación de ELI es compleja. Una visión como la que se aporta en el Marco Teórico puede facilitar la comprensión de su alcance y de las potencialidades que ofrece. Así, el resultado de este trabajo puede facilitar la definición de una hoja de ruta, para que cada sistema fije sus prioridades y avanzar en concretando las subsiguientes fases y definir sus respectivas especificaciones técnicas.

Otro objeto de estudio se refiere a los aspectos relacionados con la explotación de la información y análisis jurídico por parte de agentes fuera del sector público. Mientras que los servicios de las publicaciones oficiales son ofrecidos en exclusiva por las Administraciones Públicas, servicios como los relacionados con el análisis jurídico pueden ser prestados por el sector privado, por lo que las Administraciones

Públicas deben garantizar el acceso a sus recursos en condiciones de igualdad y de forma que se facilite el desarrollo de estos servicios en el mercado.

Por último, indicar que el campo de investigación sobre ELI, al ser un ámbito que está poco explotado en el ámbito académico y que tiene un amplio recorrido, promete ser un campo abierto a significativas líneas de investigación futuras. Además, dado que todavía en España la implantación de ELI es reducida, esto puede ser objeto de mayores investigaciones sobre el seguimiento de las implantaciones, así como del estudio y análisis de soluciones y nuevos servicios que aprovechen los beneficios de ELI para que la sociedad disponga de una mejor información jurídica.

Bibliografía

- Alvite Díez, M. L. 2009. Las bases de datos jurídicas y el uso del lenguaje XML en España. *Scire*, 15(1):33-57
- Barabucci, G., Cervone, L., Palmirani, M., Peroni, S., and Vitali, F. 2009. Multi-layer markup and ontological structures in Akoma Ntoso. In *International Workshop on AI Approaches to the Complexity of Legal Systems*. Springer, 133-149.
- Berners-Lee, T., Hendler, J., y Lassila, O. 2001. The semantic web. *Scientific American*, 284(5):34-43.
- Casanovas, P., Palmirani, M., Peroni, S., van Engers, T., y Vitali, F. 2016. Semantic web for the legal domain: the next step. *Semantic Web*, 7(3):213-227.
- Chalkidis, I., Nikolaou, C., Soursos, P. y Koubarakis, M. 2017. Modeling and Querying Greek Legislation Using Semantic Web Technologies. In *European Semantic Web Conference*. Springer, 591-606.
- Francart, T., Dann, J., Pappalardo, R., Malagon, C., Pellegrino, M. 2019. The European Legislation Identifier. *Frontiers in Artificial Intelligence and Applications*, 317:137-148.
- Francesconi, E., Küster, M. W., Gratz, P., y Thelen, S. 2015. The ontology-based approach of the publications office of the eu for document accessibility and open data services. In *International Conference on Electronic Government and the Information Systems Perspective*. Springer, 29-39.
- Orenga-Gaya, L. y Giralt, O. 2011. El diari oficial de la generalitat de Catalunya: génesis de un diario digital. *El profesional de la información*, 20(3):351-355.
- Pandit, H. J., Fatema, K., O'Sullivan, D., y Lewis, D. (2018). GDPRTEXT-GDPR as a linked data resource. In *European Semantic Web Conference*. Springer, 481-495.
- Rodríguez-Doncel, V., Navas-Loro, M., Montiel-Ponsoda, E., y Casanovas, P. 2018. Spanish legislation as linked data. In *TERECOM@ JURIX*, 135-141.

Rodríguez-Vela, C. (2011). Información de la Administración General del Estado en la era digital: Publicaciones oficiales. *El profesional de la información*, 20(3): 315–323.

van Opijnen, M., y Ivantchev, A. 2015. Implementation of ECLI–State of Play. In *Legal Knowledge and Information Systems: JURIX 2015: The Twenty-Eighth Annual Conference*, 165-168.