

# CROSS-FOREST, armonización y modelización de datos. Un proyecto transfronterizo de datos forestales abiertos de España y Portugal

REVISTA **MAPPING**  
Vol. 28, 198, 38-44  
noviembre-diciembre 2019  
ISSN: 1131-9100

*CROSS-FOREST, data harmonization and modelization.  
A cross-border project of open forest data from Spain and Portugal*

<sup>(1)</sup> Ramón Baiget Llompart, <sup>(4)</sup> Guillermo Vega Gorgojo, <sup>(1)</sup> Marta Lerner Cuzzi,  
<sup>(4)</sup> José Miguel Giménez, <sup>(1)</sup> Belén Fierro García, <sup>(2)</sup> Alexandra Fonseca,  
<sup>(2)</sup> Ana Luísa Gomes, <sup>(3)</sup> Vicente Matellán Olivera, <sup>(3)</sup> Jesús Lorenzana Campillo,  
<sup>(4)</sup> Felipe Bravo Oviedo, <sup>(4)</sup> Cristóbal Ordóñez Alonso, <sup>(1)</sup> Víctor Gonzalvo Morales,  
<sup>(1)</sup> Asunción Roldán Zamarrón

## Resumen

El proyecto Cross-Forest<sup>1</sup> pretende desarrollar infraestructuras de servicios digitales (DSIs) orientados (i) a la obtención de itinerarios selvícolas que permitan estimar la evolución de las masas forestales y la calidad de la madera a nivel de país y (ii) al control de incendios forestales a través de información precisa sobre materiales combustibles y modelos de propagación. Dichos objetivos se basan en los conjuntos de datos forestales geográficos y alfanuméricos de Portugal y España. Se emplean recursos de computación de alto rendimiento (HPC) debido a la complejidad de los modelos y a la necesidad de realizar numerosas simulaciones con distintas configuraciones. Para ello, Cross-Forest está elaborando una ontología transfronteriza de datos forestales en colaboración con las Administraciones Públicas de Portugal y España, y proporciona un repositorio público (Endpoint) basado en las especificaciones de la Directiva 2007/2/CE (INSPIRE) y de datos abiertos enlazados (LOD) para publicar los datos forestales según la ontología producida. El modelo y la publicación de datos abiertos puede ser utilizable por profesionales forestales y ciudadanía en general, que, de esta forma, pueden tener acceso completo y fácil a los datos forestales producidos por España y Portugal. Asimismo, el análisis de la propagación y evolución de incendios puede resultar de gran utilidad para los gestores implicados.

<sup>(1)</sup><https://crossforest.eu/>

## Abstract

Cross-Forest aims at developing Digital Service Infrastructures (DSIs) focused on: (i) obtaining predictions that allow to estimate the evolution and wood quality of forests at a national scale, and (ii) assessing and controlling forest fires using information related to forest fuels and propagation models. These objectives will be based on the forest datasets (geographic and alphanumeric) from Portugal and Spain. High-Performance Computing (HPC) resources will be used, due to the complexity of the models to be employed and to the need of performing multiple simulations under different configurations. With that goal, Cross-Forest is developing a cross-border ontology for forest data in collaboration with the Portuguese and Spanish Public Administrations, and will provide a public repository (Endpoint) based on the recommendations from INSPIRE and Linked Open Data (LOD), where forest data will be published, in accordance with the ontology. The data model and publication of LOD a basis for forest professionals and general citizenship, who will have easy and free access to the forest datasets produced by Portugal and Spain. Likewise, the assessment of forest fires evolution and propagation will be very useful for forest managers.

**Palabras clave:** transfronterizo, Linked Open Data, SPARQL, INSPIRE, HPC, Información geográfica forestal, incendios, modelización, ontología, RDF.

**Keywords:** cross-border, Linked Open Data, SPARQL, INSPIRE, HPC, Geographic forest information, forest fires, modelization, ontology, RDF.

<sup>(1)</sup> Grupo Tragsa

<sup>(2)</sup> Direção-Geral do Território (Portugal)

<sup>(3)</sup> Fundación Centro de Supercomputación Castilla y León (SCAYLE)

<sup>(4)</sup> Universidad de Valladolid

[rbl@tragsa.es](mailto:rbl@tragsa.es), [quiveg@tel.uva.es](mailto:quiveg@tel.uva.es)

Recepción 12/12/2019

Aprobación 16/12/2019

## 1. INTRODUCCIÓN

La demanda de recursos naturales, especialmente los bosques, combinado con un clima cambiante y amenazas tales como los incendios forestales que azotan muchas zonas en Portugal y España, plantean serios riesgos tanto para los sistemas naturales como para las sociedades humanas. Por ello, una política adecuada de gestión forestal orientada a la ciudadanía y a la preservación del medio ambiente no solo debe basarse en acciones reactivas, sino también en la prevención y la formulación de políticas basadas en modelos bien fundamentados, nutridos con datos de calidad.

En una época en la que la cantidad de datos disponibles crece rápidamente, las entidades públicas son cada vez más conscientes del valor de compartir los datos de manera coherente e interoperable. La Unión Europea (UE), de conformidad con la Directiva sobre la reutilización de la información del sector público (Directiva 2003/98/CE) y su posterior revisión (2013/37/UE), ha alentado la creación de infraestructuras de servicios digitales (DSI) en el sector público. Los DSI son plataformas temáticas que agregan información gratuita y armonizada de varios países que permite la reutilización y combinación de datos públicos abiertos para el desarrollo de productos y servicios de información para ciudadanos, empresas y administraciones públicas. El acceso mejorado a la información del sector público es un componente esencial de la estrategia de la UE para estimular nuevas oportunidades de negocio y abordar los desafíos sociales, incluida la investigación, con un enorme potencial de reutilización aún sin explotar, especialmente a nivel transfronterizo e intersectorial.

Las Administraciones Públicas hacen un gran esfuerzo para generar datos de calidad, por ejemplo, en el dominio forestal, incluyendo: inventarios, mapas, información del suelo e incendios, entre otros. Sin embargo, generalmente los datos no son fácilmente accesibles, no están integrados y presentan diferentes modelos y formatos. Estos hechos limitan en gran medida el desarrollo de servicios y aplicaciones a partir de dichos datos.

Cross-Forest tiene como objetivo abordar este problema de acceso e integración de datos forestales. Cuando las fronteras geográficas no tienen ningún sentido, como ocurre con muchas cuestiones medioambientales, la publicación común, compartida y transfronteriza de datos, utilizando tecnologías de la Web Semántica como Linked Open Data (LOD), puede ser una herramienta muy poderosa para responsables políticos, administraciones públicas, expertos forestales y ciudadanos en general.

LOD es la mejor tecnología existente para publicar datos abiertos (Bizer et al., 2009). El uso de ontologías o vocabularios comunes y la integración de datos, así como la inclusión de enlaces a fuentes de datos externas, permiten enriquecer la red de datos existente. Sin embargo, hay otras áreas donde

esta publicación de información no es suficiente y donde también es necesario desarrollar algoritmos que, utilizando LOD, puedan modelar procesos complejos que involucran grandes volúmenes de datos (por ejemplo, incendios forestales).

Cross-Forest persigue desarrollar infraestructuras de servicios digitales orientadas al control de incendios forestales y la evaluación del crecimiento forestal, mediante el uso de modelos que requerirán computación de alto rendimiento recursos (HPC) para funcionar correctamente. Los cimientos de estos servicios son los conjuntos de datos forestales geográficos de Portugal y España, utilizando el enfoque de LOD para compartir y publicar los datos georreferenciados de las administraciones públicas de ambos países.

Mediante la combinación de conjuntos de datos de inventario forestal, mapas forestales, uso del suelo / cobertura del suelo, datos de observación, topográficos y climáticos Cross-Forest desarrollará:

- (1) modelos de propagación de incendios forestales y sus efectos con el fin de proporcionar a la Administración Pública información sobre posible evolución predicción de severidad, y recomendaciones útiles sobre medidas posteriores a los eventos (Piloto FRAME).
- (2) modelos para estimar la situación actual de la calidad de la madera de las principales especies forestales de un área, así como simulaciones de cómo evolucionarían en diferentes escenarios silvícolas con diferentes intensidades de manejo (Piloto CAMBRIC).

Los conjuntos de datos abiertos generados en Cross-Forest podrán conectarse con la información LOD disponible a través de otros nodos, como el Endpoint de Cross-Nature (proyecto CEF ya finalizado), que proporciona acceso a información sobre especies exóticas invasoras y protección de la biodiversidad.

Este artículo presenta los primeros avances de Cross-Forest, es decir, una introducción a la parte de generación de datos relacionada con ontologías y LOD y los pilotos CAMBRIC y FRAME, incluyendo los recursos HPC necesarios para las simulaciones requeridas.

## 2. MATERIAL Y MÉTODO

### 2.1 Conjuntos de datos

Cross-Forest debe publicar en LOD los Inventarios Forestales de España y Portugal. Para ello, se analizan los siguientes inventarios y sus respectivas cartografías.

El **Inventario Forestal Nacional de España (IFN)** es un proyecto continuo, homogéneo para todo el territorio español, en el cual se repiten las mediciones cada 10 años, recorriéndose todo el territorio nacional en cada ciclo decenal.

Se basa en la toma de datos con parcelas de un muestreo

---

## La Ley de Montes, en su artículo 28, establece la coordinación de la Administración General del Estado para la elaboración de la información Forestal Española, que incluye, entre otros, el Inventario Forestal Nacional y el Mapa Forestal de España

---

sistemático realizado en la superficie forestal arbolada a nivel provincial. Los datos obtenidos en los trabajos de campo llevan un proceso, en gran parte informatizado, que proporciona información exhaustiva en forma de tablas, mapas y bases de datos alfanuméricas y cartográficas, disponibles tanto en las publicaciones del IFN como en los ficheros descargables a través de la web institucional. Actualmente se encuentran en descarga el Tercer IFN (1997-2007). El 4º IFN está actualmente en proceso de publicación.

Los datos recopilados se publican en una base de datos SQL, con un esquema ad-hoc y en formato Microsoft Access. Esto dificulta el uso de los datos para personas fuera del dominio forestal e incluso para personas en el dominio que no conocen la metodología específica utilizada para recopilar los datos.

El **Mapa Forestal de España (MFE)** es la cartografía básica forestal a nivel estatal, que recoge la distribución de los ecosistemas forestales españoles. Su objetivo principal es ser la cartografía base del IFN. Por ello, análogamente al IFN, la unidad básica de trabajo es la provincia y su periodicidad es decenal, recorriéndose todo el territorio nacional en cada ciclo a lo largo de una década.

Se realiza con una metodología basada en fotointerpretación sobre imagen digital donde se identifican teselas homogéneas y sus características (especies, estructura, Fracción de Cobertura Cubierta entre otras). Posteriormente, se realiza una comprobación en campo de un porcentaje de las teselas fotointerpretadas. Actualmente, se encuentran en descarga el MFE50 (escala 1:50.000) desarrollado entre 1998 y 2007, y el MFE de máxima actualidad, donde conviven provincias del MFE25 (escala 1:25.000) y MFE50 para aquellas provincias donde no se ha finalizado la versión 25.

Los datos recopilados se dividen en archivos en formato GIS, donde cada tesela está caracterizada por un polígono. Si bien estos archivos son de formato común para la comunidad SIG, no lo son para otros usuarios que no puedan interpretar fácilmente sus contenidos.

El **Inventario Forestal Nacional (IFN) portugués**, es una cartografía con metodología similar al IFN español, dispone de una fase de fotointerpretación y otra de trabajo de campo. Actualmente está en proceso de elaboración la

sexta versión del IFN (año de referencia 2015).

**Carta de Uso y Ocupación del Suelo (COS).** Es una cartografía de Portugal basada en fotointerpretación de la ocupación del suelo con nomenclatura similar al CORINE Land Cover. Está disponible mediante servicio WMS. La última versión descargable es el COS2015, la 6ª versión de COS2018 está actualmente en marcha. La gran diferencia entre los inventarios españoles y portugueses es que actualmente no hay conexión entre los dos inventarios portugueses (IFN y COS).

### 2.2. Ontología forestal y datos abiertos enlazados

El término datos abiertos enlazados (*Linked Open Data* — LOD) hace referencia a un conjunto de buenas prácticas para la publicación de datos de manera estructurada en la Web (Bizer et al., 2009). El objetivo es crear una base de datos distribuida a escala global y basada en estándares abiertos a la que suele referirse como Web de Datos o Web Semántica (Heath y Bizer, 2011). Las tecnologías semánticas en las que se basa LOD son el resultado de un largo proceso de investigación por la comunidad científica y estandarización por el W3C. Así, RDF es el lenguaje que permite crear anotaciones de datos; RDFS y OWL son lenguajes para la creación de ontologías (Allemang y Hendler, 2011), es decir, los modelos con los que se anotan los datos; y SPARQL es el lenguaje de consultas.

Uno de los objetivos de es la publicación como LOD de los datos forestales presentados en la sección 2.1. Cross-Forest Para ello, una de las tareas del proyecto es la creación de un conjunto de ontologías que conceptualicen el dominio en lo que respecta a los datos de inventario y mapa forestales. Aunque el foco de Cross-Forest está limitado a España y Portugal, el ámbito de las ontologías a desarrollar es más amplio para que puedan emplearse en otros contextos.

Para el desarrollo de las ontologías se ha partido del esquema relacional en el que los datos se encuentran actualmente. Se ha hecho uso de la documentación que describe las tablas, y se ha trabajado conjuntamente entre los ingenieros de ontologías y los expertos del dominio para entender los conceptos y relaciones existentes en los datos. En su diseño se han utilizado las mejores prácticas para la publicación de datos en la Web (Lóscio et al., 2017) y para la publicación de datos espaciales en la red (Tandy et al., 2017). Asimismo, se han utilizado patrones conocidos tanto para describir las posiciones de las entidades espaciales como las diferentes medidas, ej. área basimétrica. El diseño de la ontología se ha guiado por estos dos principios:

1. Modularidad: Se han definido 5 ontologías “core” más una serie de módulos para relacionarlas con los datos originales de las tablas y con ontologías externas.
2. Alineamiento explícito: No se reutilizan directamente los términos de otras ontologías, sino que se definen términos propios y se hace un alineamiento con ontologías y datos externos mediante los módulos anteriormente mencionados.

Con estos principios se busca maximizar la posibilidad de reutilización de los datos generados. Por un lado, permiten a un usuario utilizar únicamente las ontologías y módulos que deseen, y enlazando los datos con sus ontologías de preferencia. Por otro lado, permiten hacer reutilización segura de vocabularios externos.

Los datos del IFN se han extraído y combinado, de tal forma que todas las tablas con el mismo nombre de cada fichero (recordemos los datos publicados, tanto del IFN como del MFE, se encuentran separados en ficheros por provincia) se han unido en una única tabla, en la que existe un campo adicional para indicar la provincia de los datos. Estas tablas se han procesado utilizando *SPARQL-generate* (Lefrançois et al., 2017) para generar los datos en RDF.

Los datos del MFE (que se encuentran en ficheros *shapefile*, propios de herramientas GIS) se han preprocesado para transformarlos a *geojson*. En este preprocesado se han generado varias capas en las que los polígonos se combinan entre sí y tienen diferentes niveles de simplificación. Esto permite aumentar el rendimiento en consultas que se refieran a un territorio extenso pero no requieran gran nivel de precisión. Posteriormente, los ficheros *geojson* se han procesado con *SPARQL-generate* para generar los datos en RDF.

Adicionalmente, se propone en el proyecto la publicación de datos espaciales en LOD de tres mallas terrestres de dimensiones 25x25 metros, 1x1 kilómetro y 10x10 kilómetros. A cada una de las celdas contenidas en las mallas se le asigna un identificador único (URI) que permite la representación de los objetos espaciales de acuerdo con las recomendaciones sobre buenas prácticas en publicación de datos en la Web (Lóscio et al., 2017). De esta forma, cualquier publicación LOD con distinta capa de información, pero con una malla geográfica homogénea puede reutilizar los URIs de las celdas para hacer referencia a la localización de sus elementos.

### 2.3. Simulaciones en entorno HPC: pilotos CAMBRIC y FRAME

En Cross-Forest se está trabajando en dos pilotos o casos prácticos, que consisten en simulaciones en un entorno HPC utilizando los datos abiertos y enlazados generados a partir del MFE y del IFN. En este entorno común se emplean dos simuladores: SIMANFOR (Bravo et al 2012) en el piloto CAMBRIC y el motor de propagación de ERVIN en el piloto FRAME, que trabajan en dos entornos forestales con marcos temporales y espaciales bien diferenciados: 1) la simulación de evolución de masas forestales en condiciones estables y su consecuencia en la calidad de la madera resultante y 2) la aparición de incendios y la evolución de los mismos en función de las condiciones climáticas y los medios empleados en su extinción.

En el **piloto CAMBRIC** se van a realizar simulaciones que permitan estimar la calidad de la madera en masas puras y mixtas de las especies forestales más representativas: *Pinus*

*sylvestris*, *Pinus pinaster*, *Pinus nigra*, *Fagus sylvatica*, *Quercus pyrenaica*, *Quercus robur* y *Quercus petraea*.

El flujo de trabajo en CAMBRIC es el siguiente:

1. Selección de las áreas que incluyen las especies objetivo a partir del Mapa Forestal Español,
2. Extracción de los datos del IFN incluidos en esas áreas,
3. Estimación de la calidad de la madera en las áreas seleccionadas,
4. Simulación de las diferentes alternativas selvícolas apropiadas para cada especie y combinación encontradas
5. Publicación de escenarios más adecuados para cada zona y cuantificación de la cantidad y calidad de madera que se puede obtener en cada uno de ellos.

Este proceso está basado en los datos del IFN disponibles y en los modelos de crecimiento y producción ya existentes (Bravo et al 2011) para masas puras programados de forma que sea posible usarlos tanto en masas puras como en masas mixtas y los desarrollados con la misma estructura a partir de los datos del IFN de bosques mixtos cuando no se dispone de modelos. Se está diseñando una variable indicadora de la calidad de la madera compatible con los datos del IFN y los modelos a utilizar.

Los escenarios de simulación que se se ensayan están basados en Duncker et al. (2012) y consisten en 4 opciones bien diferenciadas:

- Escenario de no intervención, que consiste en la ausencia de gestión, con ninguna corta planificada a lo largo de la vida de la masa.
- Escenario próximo a la naturaleza, que imite la mortalidad natural que ocurriría en la masa, pero anticipándose con cortas programadas que sean compatibles con funciones ecológicas del bosque.
- Escenario de objetivos combinados, en el que se pretende satisfacer simultáneamente la explotación de madera y funciones ecológicas del bosque.
- Escenario productivo, en el que el objetivo principal es la obtención de madera. Para este escenario sólo es adecuada la especie *Pinus pinaster atlantica* en masas coetáneas y monoespecíficas.

Los resultados obtenidos de cada escenario deben ser publicados en forma de datos abiertos y enlazados utilizando la ontología forestal que se ha desarrollado en este proyecto.

En el **piloto FRAME** se realizan simulaciones de comportamiento de incendio forestal, sobre tamaños de celda de propagación de un metro cuadrado, obteniendo los valores de entrada, relativos al modelo digital de elevaciones, los viales e infraestructuras, los modelos de combustible y los datos meteorológicos, a partir de los datos publicados en abierto de las distintas administraciones.

Concretamente, se utiliza una modificación del motor de

propagación desarrollado para el entrenador virtual de incendios forestales ERVIN, sobre el entorno HPC mencionado, utilizando a su vez como interfaz de acceso la plataforma de gestión de combate y planificación EMERCARTO.

La ejecución de una simulación consta de los siguientes pasos:

1. Selección del área o parcela de simulación
2. Selección de los puntos o líneas de ignición
3. Ejecución del modelo a partir de los datos de entrada accesibles a tiempo real
4. Modificación de las condiciones y acciones de combate sobre el escenario propuesto

Ambos pilotos deben ser ejecutados de forma operativa en el superordenador Caléndula, gestionado por SCAYLE. Entre las principales características técnicas con las que cuenta Caléndula podemos destacar una capacidad de cálculo total de 388 TFlops (Rpeak) que es aportada por sus 310 servidores. Estos servidores cuentan con una capacidad agregada de memoria de 16.7 TB de memoria y en cuanto al almacenamiento disponible para cálculo científico, su capacidad asciende a 650 TB. En el proyecto se utilizan diferentes tecnologías de paralelización como OpenMP o MPI, todo ello sobre la infraestructura software de gestión de colas basada en Slurm.

### 3. RESULTADOS

Las ontologías resultantes son las siguientes:

- Simple Positions Ontology: Esta ontología permite representar posiciones de entidades, tanto absolutas (usando un sistema de coordenadas de referencia) como relativas (a otra posición).
- EPSG: Esta ontología contiene una descripción los sistemas de coordenadas de referencia. Se usa para describir posiciones absolutas en la Simple Positions Ontology.
- Simple Measures Ontology: Permite definir medidas tomadas en entidades físicas, describiendo su valor y sus unidades.
- IFN: Permite describir los datos del Inventario Forestal Nacional, incluyendo por el momento parcelas y árboles. Sus posiciones y medidas se describen utilizando las ontologías Simple Positions Ontology, EPSG, y Simple Measures Ontology.
- MFE: Permite describir los datos de Mapa Forestal Nacional. Los datos sobre especies se toman de la ontología IFN, y los datos sobre posiciones y medidas se describen utilizando las ontologías Simple Positions Ontology, EPSG, y Simple Measures Ontology.
- Spatial grid ontology: Especifica un sistema de malla geográfica para la publicación de datos espaciales de manera multipropósito.

Además de las ontologías existen dos módulos para relacionar los datos de las ontologías IFN y MFE con las tablas originales, así como una serie de módulos para enlazar los términos de las ontologías con otras ontologías y vocabularios externos. Entre las ontologías externas cabe mencionar a GeoSPARQL<sup>(1)</sup> para la publicación de datos de objetos espaciales; QUDT<sup>(2)</sup> para la definición de unidades de medida, tipo de cantidad y dimensiones; y el Vocabulario para la representación de datos de territorio<sup>(3)</sup> en España y sus unidades administrativas. Puede verse un esquema del conjunto de ontologías definido en la Figura 1.

Los datos generados suponen un total de aproximadamente 4GB y 73 millones de triplas, en los que se describen aproximadamente 97 000 parcelas y 1,4 millones de árboles del IFN, y unas 747 000 teselas de MFE. Se han cargado estos datos en un almacén de triplas que permite la ejecución de consultas en el lenguaje SPARQL. La Figura 2 muestra una captura de pantalla de una herramienta de exploración que obtiene los datos del IFN y del MFE a partir del mencionado almacén de triplas.

En el piloto CAMBRIC se espera obtener una cuantificación de las existencias actuales de madera y su calidad para las especies seleccionadas, así como su evolución en distintos escenarios silvícolas, a partir de las simulaciones que se van a ejecutar en el entorno HPC de SCAYLE.

En el piloto FRAME se espera obtener una simulación de incendio forestal, sin límite espacial, sobre la que puedan ejecutarse acciones de extinción y modificaciones al vuelo de las condiciones de partida, sobre distintos escenarios operativos, aprovechando igualmente la potencialidad del entorno HPC de supercomputación.

### 4. CONCLUSIONES

En el proyecto Cross-Forest se pretende implantar infraestructuras de servicios digitales para el desarrollo de dos pilotos entre España y Portugal: la estimación de la evolución de las masas forestales y la calidad de la madera (CAMBRIC), y el control de incendios forestales (FRAME). Estos dos pilotos requieren datos forestales de calidad, accesibles, integrados y enlazados. Para ello se ha desarrollado un conjunto de ontologías forestales con las que pueden anotarse entidades espaciales, mallas geográficas, inventarios y mapas forestales. Estas ontologías se han empleado para generar una versión LOD de los datos del IFN y del MFE. El volumen total es de unos 4GB de 73 millones de triplas en RDF que serán ofrecidas para el

<sup>(1)</sup><https://www.opengeospatial.org/standards/geosparql>

<sup>(2)</sup><https://www.qudt.org/>

<sup>(3)</sup><http://vocab.linkeddata.es/datosabiertos/def/sector-publico/territorio>



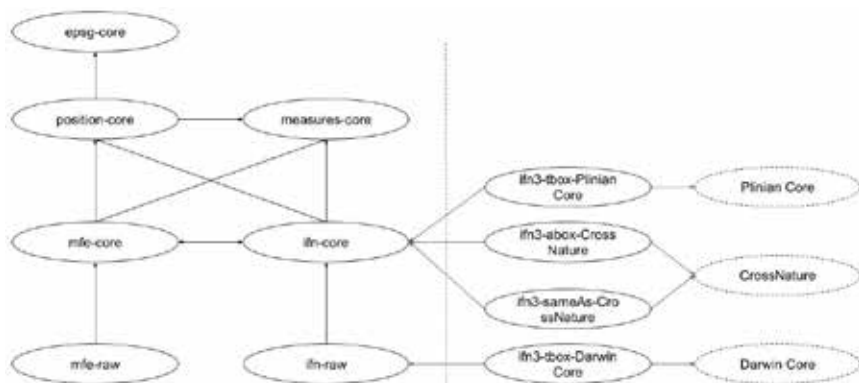


Figura 1. Ontologías Cross-Forest. A la izquierda se encuentran las ontologías y los módulos de enlace con las tablas, a la derecha se encuentran los módulos de enlace con ontologías externas



Figura 2. Captura de pantalla de una herramienta de exploración del IFN y del MFE que consume los datos del almacén de triplas preparado en Cross-Forest

público próximamente. De esta manera se proporcionará un acceso completo y fácil a la gran cantidad de datos producidos por España y Portugal para los profesionales forestales y ciudadanía en general. Asimismo, el análisis de la propagación y evolución de incendios será de gran utilidad para los gestores implicados. Como trabajo futuro se seguirá con el desarrollo de los pilotos CAMBRIC y FRAME utilizando los recursos de supercomputación disponibles en Cross-Forest.

Los beneficios esperados con el proyecto son los siguientes:

- Creación de un conjunto de ontologías que conceptualicen el dominio en lo que respecta a los datos de inventario y mapa forestal. Además, la incorporación de un sistema de URIs para los datos forestales con malla geográfica homogénea que podrá conectar con otras capas de información.
- Aplicación para herramientas de gestión forestal, modelización de la evolución de masas forestales y propagación de incendios. Los resultados obtenidos de cada escenario serán publicados en forma de datos abiertos y enlazados utilizando la ontología forestal

desarrollada.

- Utilización de recursos de supercomputación y aplicación problemas del ámbito forestal.
- Acceso completo y fácil para los profesionales forestales y ciudadanía en general, a la gran cantidad de datos producidos por España y Portugal mediante el modelo y la publicación de estos datos en formato abierto. Asimismo, el análisis de la propagación y evolución de incendios será de gran utilidad para los gestores implicados.

## AGRADECIMIENTOS

Cross-Forest (2017-EU-IA-0140) está cofinanciado por la UE a través del programa CEF (*Connecting Europe Facility*) que coordina INEA (*Innovation and Networks Executive Agency*).

Cross-Forest cuenta con apoyo institucional del Ministerio de Agricultura, Pesca y Alimentación (MAPA - SG de Política Forestal) y el Instituto da Conservação da Natureza e Florestas (ICNF) de Portugal. Ambos organismos actúan como proveedores de datos y forman parte del Comité Asesor. Se cuenta también con la participación en el Comité Asesor de investigadores de CIFOR- INIA, UPM e Idearium, expertos en las distintas temáticas del proyecto.

Para la publicación de datos geográficos en formato abierto se siguen las recomendaciones del Subgrupo de Trabajo de Datos Geográficos Abiertos Enlazados creado en 2018 en la reunión del GT-IDEE en el marco de las JIIDE 2018 celebradas en Mahón.

## REFERENCIAS

Aguirre F, Benito I, Carrillo A, Díaz M, González D, Gonzalvo V, López J., 2017. El Grupo Tragsa y los Incendios Forestales. Proyecto AF3., Integración de tecnologías de apoyo en la gestión de incendios. Revista Montes 127: 29-35.

Aguirre F, Carrillo A, Gonzalvo V, Hombrados E., 2017. ERVIN instruye de forma virtual en incendios forestales. Revista Montes 130: 37-40.

Allemang, D., Hendler, J. (2011). Semantic web for the working ontologist: effective modeling in RDFS and OWL (2nd edition). Morgan Kaufmann.

Bizer, C., Heath, T., Berners-Lee, T. (2009), Linked Data: The Story So Far, International Journal on Semantic Web and Information Systems, 5(3):1-22.

- Bravo F, Álvarez-González JG, del Río M, Barrio M, Bonet JA, Bravo-Oviedo A, Calama R, Castedo-Dorado F, Crecente-Campo F, Condés S, et al., 2011. Growth and yield models in Spain: Historical overview, contemporary examples and perspectives. *Forest Syst* 20(2): 315-328. <http://dx.doi.org/10.5424/fs/2011202-11512>
- Bravo, F.; Rodríguez, F.; Ordóñez, A.C. A web-based application to simulate alternatives for sustainable forest management: SIMANFOR. *For. Syst.* 2012, 21, 4–8.
- Dunker, P., Barreiro, S., Hengevelds, G., Lind, T., Mason, W., Ambrozy, S., & Spiecker, H. (2012). Classification of forest management approaches: A new conceptual framework and its applicability to European forestry. *Ecology and Society*, 17(4), 51.

- Heath, T., Bizer, C. (2011), *Linked Data: Evolving the Web into a Global Data Space. Synthesis Lectures on the Semantic Web: Theory and Technology*, 1:1, 1-136. Morgan & Claypool.
- Lefrançois, M., Zimmermann, A., & Bakerally, N. (2017). A SPARQL extension for generating RDF from heterogeneous formats. In *Proceedings of the Extended Semantic Web Conference, ESWC2017*. Portoroz, Slovenia.
- Lóscio, B.F., Burle, C., Calegari, N. (editors) (2017), *Data on the Web Best Practices, W3C Recommendation*, <https://www.w3.org/TR/dwbp/>
- Tandy, J., van den Brink, L., Barnaghi, P. (2017), *Spatial Data on the Web Best Practices, W3C Working Group Note*, <https://www.w3.org/TR/sdw-bp/>

## Sobre las organizaciones

### Grupo Tragsa

*El Grupo Tragsa es un conjunto de empresas públicas (TRAGSATEC y TRAGSA) integradas en la Sociedad Estatal de Participaciones Industriales (SEPI), cuya misión es proporcionar a las Administraciones soluciones integrales en cualquier tipo de necesidad ambiental, de desarrollo rural o gestión de recursos naturales, mediante un equipo humano altamente cualificado, la utilización de tecnologías innovadoras y los estándares de calidad más exigentes. El Grupo cuenta con un potente equipo especializado en cartografía y SIG aplicados a los distintos ámbitos de actuación de la empresa. Asimismo, la compañía está especializada en actuaciones relacionadas con la prevención y extinción de incendios y en la posterior recuperación de los terrenos afectados utilizando las tecnologías más punteras del sector.*

### Universidad de Valladolid (UVA)

*La Universidad de Valladolid (UVA) es uno de los centros más importantes de educación superior en España. Su equipo en el proyecto Cross-Forest está formado la Fundación UVA y los grupos de investigación iuFOR y GSIC. iuFOR (<http://sostenible.palencia.uva.es>) está formado por expertos en modelización de sistemas forestales, en simulación de alternativas de gestión sostenible y manejo y análisis de grandes bases de datos forestales; han instalado y mantienen grandes infraestructuras de recogida de datos forestales. GSIC (<https://www.gsic.uva.es/>) está formado por ingenieros y educadores expertos en telemática, Web Semántica y gestión de datos; han diseñado ontologías y publicado conjuntos de datos según los principios de datos enlazados, han desarrollado herramientas de consulta para puntos SPARQL y de exploración*

*de conjuntos de datos RDF y aplicaciones para móvil basadas en datos de DBpedia.*

### Direção-Geral do Território (DGT)

*La Direção Geral do Território (DGT) es un organismo público portugués perteneciente al Ministerio de Medio Ambiente de Portugal. La misión de la DGT es aplicar políticas públicas sobre ordenación del territorio, uso del suelo, desarrollo territorial y urbano, así como ejercer de autoridad nacional de información geográfica, estando a cargo de la red geodésica nacional, la producción de la cartografía y el Catastro, y coordinar la Infraestructura Nacional de Datos Espaciales. La DGT desarrolla investigación apoyada por proyectos nacionales o con financiación europea en distintas áreas tales como obtención del uso y ocupación del suelo partiendo de datos de satélite y otras fuentes de observación terrestre, infraestructuras de datos espaciales, el uso de información geoespacial y tecnologías en distintas áreas (por ejemplo, análisis y modelización medioambiental, planificación urbana, análisis de riesgos), participación ciudadana e información geográfica voluntaria.*

### Fundación Centro de Supercomputación Castilla y León (SCAYLE)

*SCAYLE es miembro de la Red Española de Supercomputación, infraestructura declarada por el Ministerio de Ciencia, Innovación y Universidades como ICTS Infraestructura Científica y Técnica Singular. SCAYLE tiene su sede en León (España) y proporciona recursos computacionales a todos los campos de la ciencia, gestionando diferentes clústeres informáticos, así como la red regional de investigación y educación. Su superordenador cede cerca de 30 millones de horas de cálculo a los investigadores españoles.*