

# El desafío de los identificadores persistentes y accionables

La implementación con URI basadas en el protocolo HTTP para objetos espaciales y otros recursos INSPIRE en España

**LOPEZ-PELLICER, Francisco J; BARRERA, Jesús; GONZALEZ, Julián; ZARAZAGA-SORIA, F. Javier; LOPEZ, Emilio; ABAD, Paloma; RODRÍGUEZ, Antonio F.**

Un identificador persistente es aquel que actúa como referencia estandarizada e invariante de larga duración de un recurso digital, independientemente de su estado, localización o propietario actuales. La implementación de identificadores persistentes para objetos espaciales es uno de los retos más inmediatos en la aplicación de la Directiva INSPIRE. Varios Estados miembro de la UE disponen ya de estructuras de gobernanza, procesos, normas, directrices y herramientas para generar, mantener, administrar y usar identificadores persistentes en sus respectivas IDE. Sin embargo, sus enfoques son diferentes entre sí y se encuentran en distintos niveles de madurez. INSPIRE recomienda actualmente utilizar URI basadas en el protocolo HTTP para la implementación de identificadores persistentes de objetos espaciales. No obstante, ni la gobernanza ni la efectividad en coste de la implementación de los identificadores persistentes basados en esta recomendación han sido consideradas en la Directiva INSPIRE ni en los Reglamentos que la desarrollan. Tampoco existe una estrategia de gobernanza compartida por los Estados miembros y la Comisión Europea. De hecho, desde INSPIRE se considera que lo ideal sería que cada Estado miembro estableciera autónomamente su mecanismo nacional. Este trabajo presenta los avances realizados hasta ahora en el desarrollo de una solución a ese reto.

## PALABRAS CLAVE

PID, Identificadores persistentes, NTI, ENI, HTTP URI

## INTRODUCCIÓN

Este trabajo presenta los avances realizados hasta ahora en el desarrollo de una solución al reto que suponen las URI basadas en el protocolo HTTP que, teniendo en cuenta los condicionantes ya citados, sea efectiva en coste. Esta solución debe ser capaz de registrar esas URI como identificadores persistentes de objetos espaciales INSPIRE y más adelante extender ese registro a otros recursos INSPIRE, como los conjuntos de datos espaciales y las series. Además, la solución debe ser capaz de resolver dichas URI en peticiones a servicios de descarga INSPIRE capaces de devolver una representación del recurso correspondiente en forma de fichero. En otras palabras, esta solución permitirá la implementación en España de un sistema de identificadores persistentes y accionables para objetos espaciales y otros recursos INSPIRE.

Este trabajo se organiza de la siguiente forma. Primero se presenta en detalle qué es un identificador persistente accionable. A continuación se describen los tres sistemas de identificadores persistentes accionables que más han influido en el diseño de la solución (el sistema Handle, el sistema DOI y el sistema PURL) y las implicaciones que tiene un identificador persistente accionable. Veremos a continuación cómo son los identificadores persistentes definidos por INSPIRE, por qué se pueden considerar como identificadores persistentes accionables y la propuesta que ha planteado INSPIRE para su implementación mediante HTTP URI, pero sin definir un marco europeo de referencia. La ausencia de dicho marco hace que, como se verá a continuación, cualquier diseño está condicionado por las normas nacionales (Esquema Nacional de Interoperabilidad en el caso de España) y por la diversidad y estructura de los productores de datos. Teniendo en cuenta todos esos hechos y la experiencia de las soluciones con más éxito proponemos una solución práctica. Esa solución práctica se está implementando actualmente como un prototipo que se presenta al final de este trabajo.

## ¿QUÉ ES UN IDENTIFICADOR PERSISTENTE ACCIONABLE?

Un **identificador persistente** o PID (*Persistent Identifier*) es un identificador que actúa como

referencia estandarizada e invariante de larga duración a un recurso digital independientemente de su localización física o propiedad actual. Es decir, un PID identifica de forma no ambigua un recurso digital y sigue siendo válido aun cuando cambie la localización o el dueño del recurso digital. El uso de un protocolo como HTTP para su implementación implica que además es un **identificador persistente accionable**, es decir, que el PID puede resolverse en un sistema de resolución y, en función de lo que desee el usuario y la capacidad del sistema de resolución, debe devolver si el recurso está activo:

- Una representación del recurso digital identificado o la localización de dicha representación. Por ejemplo, si el recurso digital identificado es un lugar con un nombre propio como Barcelona puede devolver una redirección a una pregunta *GetFeatureById* de un servicio de descarga INSPIRE basado en la especificación OGC WFS.
- Una representación del registro de metadatos que describe el recurso digital identificado o la localización de dichos metadatos. Por ejemplo, si el recurso digital identificado es un conjunto de datos como el Nomenclátor Geográfico Básico de España, puede devolver una redirección a una pregunta *GetRecordById* de un servicio de localización INSPIRE basado en la especificación OGC CSW.

En el caso de que el recurso hubiera sido borrado, el sistema de resolución debe informar de ello con independencia del tiempo que hubiera transcurrido desde su borrado, proporcionando metainformación en la respuesta. Por ejemplo, el sistema de PID accionables PURL que mencionaremos a continuación, utiliza la respuesta «HTTP 410 Gone» con ese propósito. Además, desde la creación del PID en adelante y mientras perviva el sistema de resolución, el sistema debe ser capaz de informar cuándo, cómo y por quién fue creado el PID, su estado actual, a dónde se resuelven tanto el recurso digital como sus metadatos, así como el histórico de cambios de propiedades esenciales, como es el estado del recurso identificado o la propiedad del PID, entre otros (véase la Figura 1).

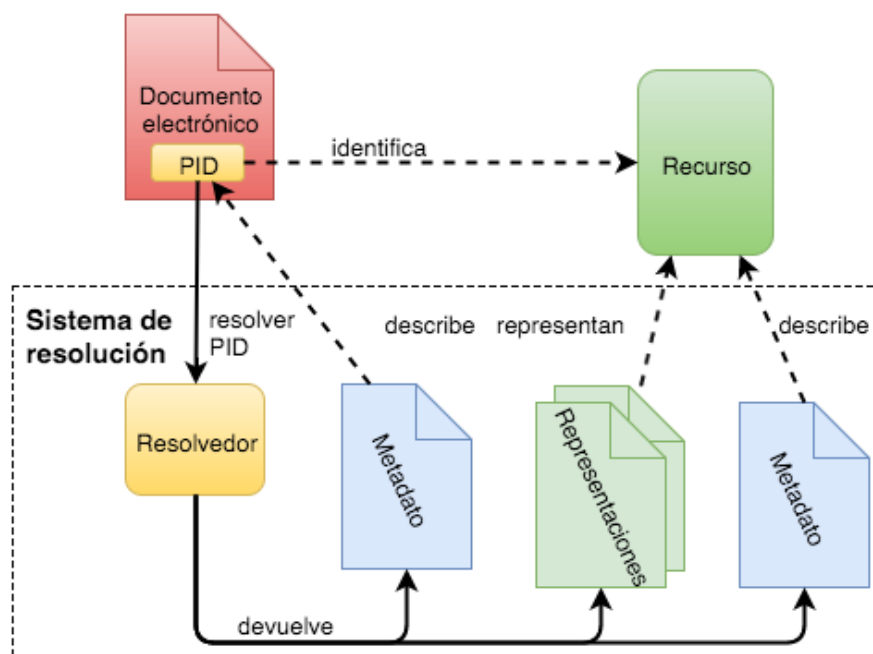


Figura 1. Elementos principales de un sistema de PID

## PRINCIPALES SISTEMAS

Los primeros sistemas de PID accionables surgen a mediados de los 90 poco después del nacimiento de la web como consecuencia de la falta de persistencia de los *Uniform Resource Locators* (URL), el nombre popular que damos a las URI que identifican recursos en la web. Una *Uniform Resource Identifiers* (URI) es una cadena de caracteres utilizada para identificar y localizar un recurso en un espacio de nombres registrados que sigue la sintaxis definida por la organización *Internet*

*Engineering Task Force* (IETF). Las URI se formalizaron por primera vez en el RFC 1630 (1994) [1]. Actualmente el estándar de referencia es el RFC 3986 (2005) [2] que define una URI como:

nombre-de-esquema «:» parte-específica [ «?» pregunta ] [ «#» fragmento ]

Donde [ ] indica una parte opcional. El nombre URL designa a las URI que identifican recursos que pueden ser accedido o localizado mediante algún protocolo (como RFC 1630 y otros), generalmente HTTP. Las buenas prácticas para su diseño están recogidas en el RFC 7320 [3] publicado en 2014.

Hay muchas iniciativas que han abordado el problema de la persistencia, pero solo vamos a centrarnos en tres atendiendo a su impacto en los últimos 25 años, ya que una de las claves de los PID es que deben ser capaces de tener una vida digital muy prolongada.

La primera solución con éxito a ese problema fue crear un sistema completamente alternativo a las URI pero que funcionara como las URI sobre Internet: el sistema Handle (1994) [4]. Este es un sistema de PID accionables distribuidos desarrollado por Bob Kahn para la *Corporation for National Research Initiatives* (CNRI) con el apoyo de *Defense Advanced Research Projects Agency* (DARPA) a principios de los 90 como alternativa a otras soluciones de la época (DNS y URL). Una descripción informativa de esa solución viene recogida en el RFC 3650 [5]. El sistema Handle incluye un sistema de resolución global con servidores locales y caché. Además, existe un registro global de autoridades que asegura la unicidad de los nombres y el enrutamiento de peticiones para gestión la resolución de los PID. El esquema de un identificador en Handle es:

autoridad «/» nombre-local

A partir del sistema Handle se desarrollaron los DOI (*Digital Object Identifier, 1997*) [6]. Los DOI son un tipo PID accionables desarrollados sobre el sistema Handle. En este sistema se almacenan junto al DOI metadatos sobre el recurso identificado que pueden contener una URI con la localización del recurso. Estos metadatos se basan en el *framework* de metadatos <indec> (*interoperability of data in e-commerce systems*), financiado por la Comunidad Europea entre 1998 y 2000, para modelar y almacenar metadatos relacionados con la propiedad intelectual. Actualmente la norma ISO 26324:2012 [7] especifica la sintaxis, el modelo y los componentes funcionales de la resolución del sistema DOI. El sistema DOI está gestionado por la fundación internacional *International DOI Foundation* (IDF). IDF además es responsable de garantizar la existencia de un sistema de resolución de DOI sobre el protocolo HTTP. El sistema DOI actualmente gestiona 114 millones de PID para más de 15 000 autoridades [8]. La emisión de un DOI puede ser gratuita si así lo decide el operador local del sistema. Un ejemplo es DataCite [9], que DOI gratuitamente a colecciones de datos. El esquema DOI es un espacio de nombres dentro del sistema Handle con la siguiente forma:

«10.» autoridad «/» nombre-local

Pero para todos es obvio que el protocolo HTTP es la forma habitual de intercambiar información hoy en día. Por ello, IDF mantiene un esquema de PID que funcionan tanto con HTTP como con HTTPS que da acceso a los PID del sistema Handle con la siguiente forma:

esquema «://» [ «dx.» ] «doi.org/10.» autoridad «/» nombre-local

Donde esquema es http o https. Cuando un PID accionable (DOI en forma HTTP) se deriva de otro PID (DOI en forma Handle) y da acceso al contenido identificado por dicho PID, se le denomina comúnmente como PID proxy.

La solución alternativa a un protocolo diferente de HTTP es establecer buenas prácticas para crear PID accionables sobre HTTP. El sistema de PURL (*Persistent URL, 1995*) [10] es el origen de la mayoría de las buenas prácticas actuales. Un PURL es un PID con forma de URI que describe una localización que se resuelve en la localización efectiva del recurso. Una de sus características es que fueron diseñados para ser persistentes. Los PURL inicialmente fueron desarrollados y mantenidos por la organización *Online Computer Library Center* (OCLC) desde mediados de los años 90 para simplificar la gestión de los catálogos que describen recursos de Internet ante la falta de progreso en

el desarrollo de un sistema persistente de nombrado de tales recursos. Otras organizaciones han montado sus propios PURL por lo que no se puede hablar de un único sistema sino de una buena práctica adoptada por diferentes organizaciones para resolver el problema de persistencia de las URI. La estructura de un PURL es:

«http://» autoridad «/» dominio «/» nombre PURL,

donde *dominio* y *nombre PURL* constituyen juntos el PURL. Además, el servidor de PURL proporciona metainformación sobre el PURL utilizando el estado de la respuesta HTTP [11].

Desde finales de 2015 el esquema original mantenido por OCLC ya no permite registrar más PURL por falta de mantenimiento [12]. Es decir, la persistencia de los PURL ha durado solo 20 años. A diferencia del sistema Handle, nunca hubo un modelo de financiación que repartiera, como es el caso del sistema Handle, el coste entre sus usuarios.

## CARACTERIZACIÓN

A partir del análisis de los sistemas de PID accionables con éxito podemos identificar una serie de propiedades comunes:

- **Consistencia verificable soportada por políticas.** El PID siempre identifica el mismo recurso. Es decir, no es reutilizable ni siquiera cuando el recurso identificado se borra o se destruye. Además, por ser accionable implica que los metadatos del PID se actualizan cuando la dirección del recurso digital se modifica o se produce algún cambio de estado relevante. Es decir, la organización que mantiene el recurso digital tiene políticas que garantizan la consistencia del PID. No solo eso, el usuario del PID puede verificar la consistencia analizando la historia, la ubicación y cualquier otra información relevante que se encuentre en los metadatos del PID.
- **Persistencia a largo plazo.** El PID tiene un periodo de validez mayor que el del recurso identificado. Además, por ser accionable implica que los metadatos del PID contienen información acerca del ciclo de vida del recurso identificado y que los sistemas de resolución utilizan esta información durante la resolución del PID para notificar el borrado del recurso, si es que se ha producido. La persistencia a largo plazo implica la necesidad de un nodo de resolución de referencia. Esta característica es esencial ya que los recursos identificados en algún momento pueden moverse de una organización a otra, la entidad que los posee puede llegar a ser disuelta, y en esos casos, o cuando los recursos se transfieren a otra organización, es imposible que dicha organización original siga resolviendo esos PID.
- **Externo y opaco.** El PID se ha diseñado para ser usado por terceros, por lo que pueden ser diferentes de los identificadores que el productor utiliza internamente. Además, por ser accionable implica que existe un mecanismo que asegura la correspondencia entre el PID y el identificador del productor, aun cuando este último cambie. No solo eso, el sistema debe proporcionar un mecanismo que asegure dicha correspondencia incluso cuando cambie el productor. Esto implica entre otras cosas, que el identificador no debe contener un nombre asociado a un productor específico.
- **Globalmente único y autorizado.** El PID accionable se ha diseñado de tal manera que existe un proceso mediante el cual se puede asegurar que no se pueden crear dos identificadores iguales. Además, por ser accionable implica necesariamente un registro centralizado de autoridades que asocie parte del espacio de PID posibles a diferentes autoridades para que puedan crear identificadores simultáneamente. Sólo deben tener derecho a crear un PID las autoridades que dan acceso digital al recurso. En caso contrario, existe el riesgo de una proliferación de PID duplicados o sin responsable.
- **Efectivo en coste.** Derivada de las dos anteriores. El esquema de PID accionables debe ser financieramente soportable. Por ejemplo, la inversión inicial en el nodo de resolución de referencia y en el registro de autoridades puede recuperarse mediante cuotas que los publicadores vean aceptables. Además, hay una economía de escala. Es decir, el coste

marginal medio de cada nuevo PID decrece a medida que aumenta el número de publicadores.

## LA DIRECTIVA INSPIRE Y LOS IDENTIFICADORES PERSISTENTES

La primera referencia en el marco INSPIRE al concepto de PID se encuentra de forma explícita en el artículo 8(2) de la propia Directiva INSPIRE. Este artículo viene a decir que los reglamentos que definen las Normas de ejecución para los conjuntos de datos que correspondan a temas del anexo I o II incluirán «*un marco común de identificación única de los objetos espaciales que sirva de referencia para situar los identificadores en los sistemas nacionales a efectos de garantizar la interoperabilidad entre ellos*». El propósito de ese marco es:

- Identificar o localizar objetos espaciales.
- Gestionar el ciclo de vida de los objetos espaciales.
- Permitir la reutilización proporcionando un identificador único a cada recurso.
- Establecer un marco de interoperabilidad entre los diferentes sistemas nacionales para identificar objetos espaciales.

Los temas afectados inicialmente son los recogidos en los anexos I y II (véase la Tabla 1).

Tabla 1. Temas afectados inicialmente por el marco de identificación única

Anexo I	Anexo II
Sistemas de coordenadas de referencia	Elevaciones
Sistema de cuadrículas geográficas	Cubierta terrestre
Nombres geográficos	Ortoimágenes
Unidades administrativas	Geología
Direcciones	
Parcelas catastrales	
Redes de transporte	
Hidrografía	
Lugares protegidos	

El Reglamento 1089/2010 que aplica la Directiva INSPIRE en lo que se refiere a la interoperabilidad de los conjuntos y los servicios de datos espaciales [13] define en sus artículos 9 y 10 que el esquema de PID para objetos espaciales requerido por la Directiva INSPIRE se regirá por los siguientes principios (véase la Tabla 2):

- Se define un tipo de datos *Identifier* para los identificadores de objetos espaciales referenciables por terceros que serán denominados «*identificador externo de objeto*» (*external object identifier*) (artículo 9(1)). Este identificador no identifica el recurso del mundo real sino que identifica el registro que representa dicho recurso en INSPIRE. El tipo de datos *Identifier* tendrá tres atributos: *localId*, *namespace* y *versionId*.
- El identificador externo de objeto identificará de manera única el objeto espacial y no se modificará durante el ciclo de vida del objeto espacial (artículo 9(2)).
- Las diferentes versiones del mismo objeto espacial serán siempre instancias del mismo tipo de objeto espacial y permanecerán invariantes *namespace* y *localId*. La duración de una versión será una propiedad del objeto espacial y estará marcada por las propiedades opcionales *beginLifespanVersion* y *endLifespanVersion* (artículo 10).

Tabla 2. Reglamento 1089/2010: Atributos del tipo de datos *Identifier*

Atributo	Definición	Tipo	Voidability
<b>localId</b>	Identificador local, asignado por el proveedor de datos. El identificador local es único dentro del espacio de nombres, es decir que ningún otro objeto espacial tiene el mismo identificador único.	CharacterString	
<b>namespace</b>	Espacio de nombres que identifica de manera única la fuente de datos del objeto espacial.	CharacterString	
<b>versionId</b>	Identificador de la versión particular del objeto espacial, con una longitud máxima de 25 caracteres. Si la especificación de un tipo de objeto espacial con un identificador externo de objeto incluye información sobre el ciclo de vida, el identificador de la versión se utiliza para distinguir entre las diferentes versiones de un objeto espacial. Dentro del conjunto de todas las versiones de un objeto espacial, el identificador de la versión es único.	CharacterString	voidable

Implícitamente ese esquema de PID para objetos espaciales y sus versiones señala la necesidad de un segundo sistema de PID que identifique la fuente de los datos. Ese segundo esquema de PID podría ser compatible con el esquema definido para los espacios de nombres en el Reglamento 1205/2008 que ejecuta la Directiva INSPIRE en lo que se refiere a metadatos [14].

En los modelos INSPIRE los atributos que utilizan el tipo de datos *Identifier* se denominan habitualmente *inspireId*. A continuación, se recogen en las Tablas 3, 4 y 5 los diferentes elementos relacionados con conjuntos de datos y series de conjuntos de datos que requieren un identificador con características de identificador persistente.

Finalmente, el Reglamento 976/2009 que ejecuta la Directiva INSPIRE en lo que se refiere a los servicios de red [15] señala que los servicios de localización deben permitir las búsquedas de metadatos mediante el identificador único del recurso. Además, los servicios de visualización deben ofrecer, en la respuesta a la petición obtener metadatos del servicio, el identificador único del recurso utilizado para crear la capa. Los servicios de descarga deben permitir utilizar el identificador del conjunto de datos espaciales para obtener dicho conjunto o su descripción. El identificador del conjunto de datos espaciales junto con el identificador único del objeto espacial será utilizado para recuperar objetos espaciales.

Analizando el marco anterior podemos deducir que:

- La implementación del Reglamento 1089/2010 requiere que se puedan registrar PID de espacios de nombres, PID de objetos espaciales y PID de versiones de objetos espaciales, y que esos PID tienen una estructura jerárquica.
- La estructura definida en el Reglamento 1089/2010 implica una centralización o coordinación del sistema de PID de espacios de nombres. La asignación de PID de objetos espaciales y de PID de versiones de objetos espaciales puede descentralizarse y delegarse en el productor, una vez fijados los espacios de nombres en los que puede crear dichos PID.
- El Reglamento 976/2009 convierte implícitamente el PID de un objeto especial y el PID de una versión de un objeto espacial en PID accionables, pero sin un servicio de resolución transparente. Es decir, no se define un mecanismo simple por el cual a partir del PID de un objeto espacial se pueda descubrir el servicio de descarga que resuelve dicho PID en una representación.

Tabla 3. Reglamento 1089/2010: Modelos que requieren PID

Modelo	Entidad	Id
Red	Elemento de red ( <i>NetworkElement</i> )	<i>inspireId</i>
	Propiedad de red ( <i>NetworkProperty</i> )	<i>inspireId</i>
Observaciones	Conjunto de observaciones ( <i>ObservationSet</i> )	<i>inspireId</i>
Complejo de actividad	Complejo de actividad ( <i>ActivityComplex</i> )	<i>thematicId</i> <i>inspireId</i>
	Permiso ( <i>Permission</i> )	<i>thematicId</i>

Tabla 4. Reglamento 1089/2010: Elementos del Anexo I que requieren PID

Tema	Entidad	Id.
Nombres geográficos	Lugar nombrado ( <i>NamedPlace</i> )	<i>inspireId</i>
Unidades administrativas	Límite administrativo ( <i>AdministrativeBoundary</i> )	<i>inspireId</i>
	Unidad administrativa ( <i>AdministrativeUnit</i> )	<i>inspireId</i>
	Condominio ( <i>Condominium</i> )	<i>inspireId</i>
	Línea de base ( <i>Baseline</i> )	<i>inspireId</i>
	Límite marítimo ( <i>MaritimeBoundary</i> )	<i>inspireId</i>
	Zona marítima ( <i>MaritimeZone</i> )	<i>inspireId</i>
Direcciones	Dirección ( <i>Address</i> )	<i>inspireId</i>
	Componente de la dirección ( <i>AddressComponent</i> )	<i>inspireId</i>
Parcelas catastrales	Unidad de propiedad básica ( <i>BasicPropertyUnit</i> )	<i>inspireId</i>
	Límite catastral ( <i>CadastralBoundary</i> )	<i>inspireId</i>
	Parcela catastral ( <i>CadastralParcel</i> )	<i>inspireId</i>
	Zonificación catastral ( <i>CadastralZoning</i> )	<i>inspireId</i>
Red de transporte	Red de transporte ( <i>TransportNetwork</i> )	<i>inspireId</i>
Hidrografía	Objeto hidrográfico ( <i>HydroObject</i> )	<i>hydroId</i>
	Cuenca de captación ( <i>DrainageBasin</i> )	<i>inspireId</i>
	Punto de interés hidrográfico ( <i>HydroPointOfInterest</i> )	<i>inspireId</i>
	Frontera tierra-agua ( <i>LandWaterBoundary</i> )	<i>inspireId</i>
	Objeto artificial ( <i>ManMadeObject</i> )	<i>inspireId</i>
	Costa ( <i>Shore</i> )	<i>inspireId</i>
	Aguas superficiales ( <i>SurfaceWater</i> )	<i>inspireId</i>
	Humedal ( <i>Wetland</i> )	<i>inspireId</i>
Lugares protegidos	Lugar protegido ( <i>ProtectedSite</i> )	<i>inspireId</i>

Tabla 5. Reglamento 1089/2010: Elementos del Anexo II que requieren PID

Tema	Entidad	Id.
Elevaciones	Cobertura de malla de elevaciones ( <i>ElevationGridCoverage</i> )	<i>inspireId</i>
	TIN de elevaciones ( <i>ElevationTIN</i> )	<i>inspireId</i>
Cubierta Terrestre	Nomenclatura de la cubierta terrestre ( <i>LandCoverNomenclature</i> )	<i>inspireId</i>
	Conjunto de datos de la cubierta terrestre ( <i>LandCoverDataset</i> )	<i>inspireId</i>
	Unidad de cubierta terrestre ( <i>LandCoverUnit</i> )	<i>inspireId</i>
	Cobertura en malla de la cubierta terrestre ( <i>RectifiedGridCoverage</i> )	<i>inspireId</i>
Ortoimágenes	Cobertura de ortoimágenes ( <i>OrthoimageCoverage</i> )	<i>inspireId</i>
	Elemento de mosaico ( <i>MosaicElement</i> )	<i>inspireId</i>
Geología	Sondeo ( <i>Borehole</i> )	<i>inspireId</i>
	Colección geológica ( <i>GeologicCollection</i> )	<i>inspireId</i>
	Objeto geológico ( <i>GeologicFeature</i> )	<i>inspireId</i>
	Objeto geofísico ( <i>GeophObject</i> )	<i>inspireId</i>
	Conjunto de datos geofísicos ( <i>GeophObjectSet</i> )	<i>inspireId</i>
	Masa de agua subterránea ( <i>GroundWaterBody</i> )	<i>inspireId</i>
	Objeto hidrogeológico ( <i>HydrogeologicalObject</i> )	<i>inspireId</i>

## IMPLEMENTACIÓN CON HTTP URI

La implementación de INSPIRE ha apostado decididamente por utilizar la web. En la web, las HTTP URI son el mecanismo utilizado para referenciar recursos de información. No solo eso, las iniciativas de interoperabilidad para las administraciones públicas de la UE apuestan por el uso de PID en forma de HTTP URI aun cuando no hay directrices ni buenas practicas establecidas [16]. En el caso de INSPIRE se recomienda su utilización, pero no va a existir un marco a nivel europeo. Literalmente, en la página de *Frequent Asked Questions* sobre la implementación de PID usado URI en INSPIRE [17], tras recomendar su uso, se señala que es innecesario ya que «ya existe, ya que INSPIRE está reutilizando recursos web existentes». Es una afirmación muy arriesgada ya que a falta de una directriz europea cada país va a seguir las normas nacionales que condicionan el diseño de un esquema de PID basado en HTTP URI.

En el caso de España estas normas son el Esquema Nacional de Interoperabilidad (ENI) [18], el cual se encuentra alineado con la Estrategia Europea de Interoperabilidad (*European Interoperability Strategy*), el Marco Europeo de Interoperabilidad (*European Interoperability Framework*) y las Normas Técnicas de Interoperabilidad (NTI) que lo desarrollan. Tal como recoge el Real Decreto 4/2010 que regula el ENI en su artículo 3 «el Esquema Nacional de Interoperabilidad y sus normas de desarrollo prevalecerán sobre cualquier otro criterio en materia de política de interoperabilidad en la utilización de medios electrónicos para el acceso de los ciudadanos a los servicios públicos». En particular la NTI de reutilización de recursos de información [19] establece cómo deben ser las HTTP URI y HTTPS URI para garantizar el direccionamiento y resolución de cualquier recurso de información reutilizable. Dado que uno de los objetivos de los PID de INSPIRE es permitir la reutilización de la información, los objetos espaciales y sus versiones quedan bajo el ámbito de esta NTI. En consecuencia, los PID deben estar publicados en una forma compatible con las reglas de esa NTI. Por suerte, esa normativa es bastante flexible y permite adaptar las URI a las necesidades de cada organización. Sólo exige para recursos de información que al menos cumplan esta estructura:

«http://» base «/recurso/» nombre-local

La base de los URI incluirá información básica sobre la procedencia de la información identificando el organismo publicador. Además establece unos requisitos básicos de comportamiento, que definen esos identificadores como PID accionables que deben utilizar los códigos de estado HTTP de forma consistente. Por ejemplo, en el caso de que sea necesario cambiar o eliminar el recurso al que apunta un identificador, se deberá establecer un mecanismo de información sobre el estado del recurso usando los códigos de estado de HTTP. En el caso de poder ofrecer una redirección a la nueva ubicación del recurso, se utilizarán los códigos de estado HTTP 3XX, mientras que para indicar que un recurso ha desaparecido permanentemente, se utilizará el código de estado HTTP 410.

Otro factor condicionante es el elevado grado de dispersión de la producción de información geoespacial derivada tanto de la estructura de la Administración General de Estado como de la descentralización territorial. No se puede asumir que la base va a ser estable ni única. Una solución óptima sería aquella que permitiera que los productores, si lo desean, tengan sus propias bases, que puedan transferir entre ellos la propiedad de los recursos y que, al mismo tiempo, garantizara la persistencia y la resolución de los PID de INSPIRE en una base bien conocida (por ejemplo <http://www.idee.es>).

Condicionantes normativos y organizativos similares de carácter nacional pueden encontrarse actualmente en otros Estados miembros de la Unión Europea.

## UNA SOLUCIÓN PRÁCTICA

La falta de una solución unificada a nivel europeo, la naturaleza descentralizada de la administración pública y la necesidad de cumplir con el ENI han impulsado el diseño de la solución que se propone a continuación. Con ella, se busca responder a los retos anteriormente identificados y es además compatible con los objetivos particulares de los diferentes actores.



La solución tiene cuatro tipos de componentes:

- **Registro de espacios de nombres.** Este registro gestiona los espacios de nombres, asociándolos a un tipo de dato, a un proveedor de datos y a uno o varios servicios de descarga INSPIRE de dicho proveedor que dan acceso a objetos espaciales de dicho tipo.
- **Registros locales de objetos espaciales y de versiones de objetos espaciales.** Los proveedores de objetos espaciales pueden crear y mantener sus propios registros locales que registrarán información de los objetos espaciales que pertenezcan a espacios de nombres que hayan registrado como propios en el registro central.
- **Registro de respaldo de objetos espaciales y de versiones de objetos espaciales.** Este registro utiliza la información del registro central del espacio de nombres para recolectar de forma automática y periódica los identificadores INSPIRE registrados en los registros locales, y, si estos no existen, directamente de los servicios de descarga. El objetivo de este registro central de respaldo es asegurar la persistencia y acceso a largo plazo de los identificadores y de la metainformación asociada.
- **Servicio de resolución.** Este servicio utiliza la información del registro de objetos espaciales para publicar bajo una misma base (por ejemplo <http://idee.es/>) todos los PID de objetos espaciales y de versiones de objetos espaciales. A todos los efectos sería esta HTTP URI sería un PID accionable:

«<http://>» base «[/recurso/](#)» namespace «[/](#)» localId [ «[/](#)» versionId ]

Este servicio resuelve las URI en redirecciones a los servicios de descarga de los publicadores. También da acceso a los metadatos que existan sobre el identificador PID.

El nodo central del sistema estaría compuesto por el registro de espacios de nombres, el registro de respaldo de objetos espaciales y de versiones de objetos espaciales y de un servicio de resolución. Los productores de datos que así lo desearan podrían tener su propio nodo compuesto de un registro local de objetos espaciales y de versiones de objetos espaciales y, opcionalmente, de un servicio de resolución. La Figura 2 muestra una representación conceptual del sistema.

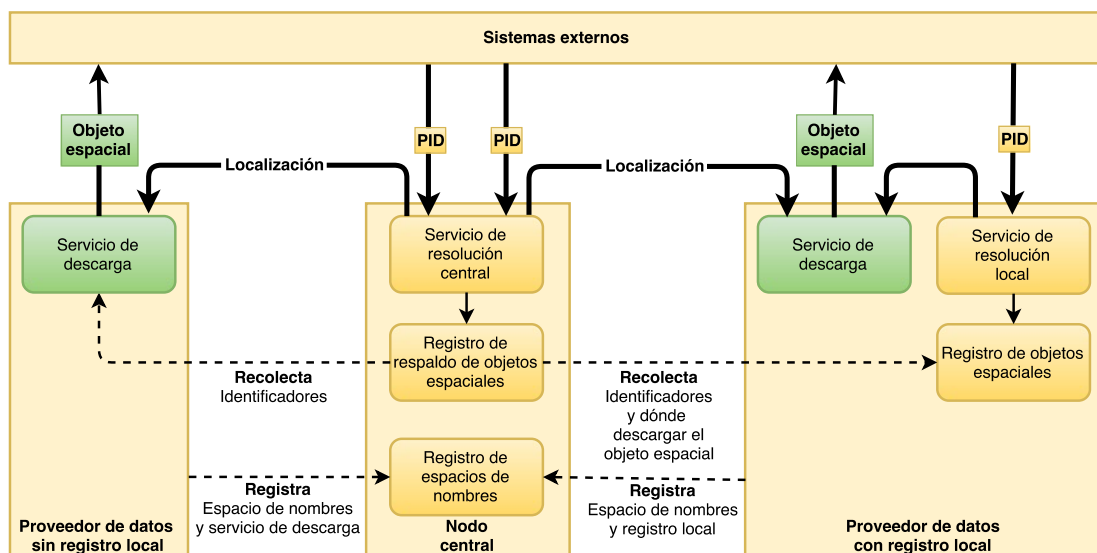


Figura 2. Arquitectura de la infraestructura

Actualmente se está desarrollando un prototipo de nodo central que tiene ya implementado un registro básico de espacios de nombres, un registro de respaldo de objetos espaciales y de versiones de objetos espaciales, y un servicio de resolución. La Figura 3 muestra diferentes capturas de pantalla de la interfaz web de los registros implementados.



Figura 2a Registro de un espacio de nombres

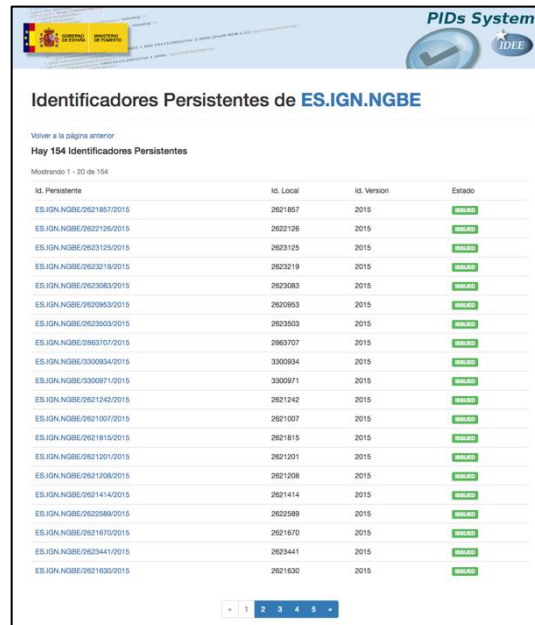


Figura 2b Listado de PID de un espacio de nombres



Figura 2c Lista de tareas de recolección de PID realizadas con éxito

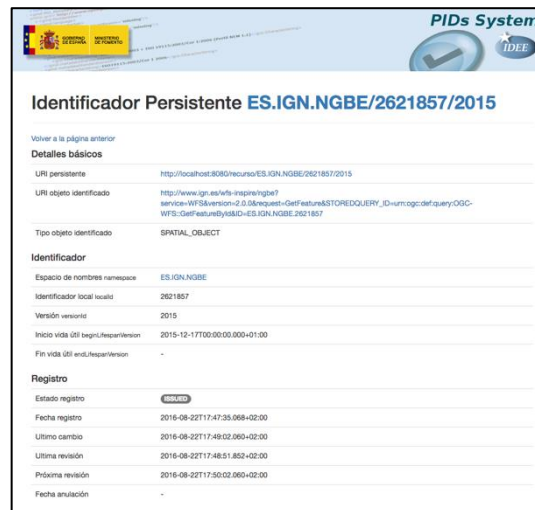


Figura 2d Metainformación sobre un PID

Figura 3 Capturas de pantalla de las interfaces web de los registros

## CONCLUSIONES

Las pruebas actuales del prototipo nos hacen confiar en la viabilidad de la solución propuesta. Sin embargo, el principal reto de un sistema de PID accionables en HTTP URI no es técnico, sino organizativo y no reside en el nodo central sino en los productores. No existe en España una NTI que defina cómo debe ser la política de identificadores persistentes de una organización, no solo para su reuso externo, sino para su consumo interno o para el cumplimiento de sus políticas. Un PID de INSPIRE es un PID INSPIRE solo si el eslabón más débil, el *localId*, el *versionId*, tienen las mismas propiedades que un PID dentro de la organización que los crea.

## AGRADECIMIENTOS

Este trabajo ha sido parcialmente financiado por el Centro Nacional de Información Geográfica (CNIG) y GeoSpatiumLab S.L.

## REFERENCIAS

- [1] T. Berners-Lee, "Universal Resource Identifiers in WWW: A Unifying Syntax for the Expression of Names and Addresses of Objects on the Network as used in the World-Wide Web," RFC 1630, Jun. 1994.
- [2] T. Berners-Lee, R. Fielding, and L. Masinter, "Uniform Resource Identifier (URI): Generic Syntax," STD 66, RFC 3986, Jan. 2005.
- [3] M. Nottingham, "URI Design and Ownership," BPC 190, RFC 7320, Jul. 2014.
- [4] L. Lannom, "Handle System Overview.," presented at the IFLA Council and General Conference, 2000.
- [5] S. Sun, L. Lannom, and B. Boesch, "Handle System Overview," Informational, RFC 3650, Nov. 2003.
- [6] N. Paskin, "Digital Object Identifier (DOI®) System ,," in *Encyclopedia of Library and Information Sciences*, 0-www.doi.org.library.touro.edu, 2010.
- [7] "Information and documentation -- Digital object identifier system," ISO 26324:2012, 2012.
- [8] "Key Facts on Digital Object Identifier System," *International DOI Foundation*, 2015. [Online]. Available: <http://www.doi.org/factsheets/DOIKeyFacts.html>. [Accessed: 13-Dec-2015].
- [9] J. Brase, "DataCite - A Global Registration Agency for Research Data," presented at the 2009 Fourth International Conference on Cooperation and Promotion of Information Resources in Science and Technology (COINFO), pp. 257-261.
- [10] "PURL," *purl.org*. [Online]. Available: <https://purl.org/>. [Accessed: 03-Nov-2015].
- [11] "PURL Help," *purl.org*. [Online]. Available: <https://purl.org/docs/help.html>. [Accessed: 08-Aug-2016].
- [12] Persistenturls Google Group, "Log in disabled," *groups.google.com*. [Online]. Available: <https://groups.google.com/forum/#!topic/persistenturls/Zpd4BHQxxIM>. [Accessed: 2015].
- [13] Unión Europea, Comisión, *REGLAMENTO (UE) No 1089/2010 DE LA COMISIÓN de 23 de noviembre de 2010 por el que se aplica la Directiva 2007/2/CE del Parlamento Europeo y del Consejo en lo que se refiere a la interoperabilidad de los conjuntos y los servicios de datos espaciales*. 2010.
- [14] European Commission, "Commission Regulation (EC) No 1205/2008 of 3 December 2008 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards metadata (Text with EEA relevance)," 2008.
- [15] *REGLAMENTO (CE) No 976/2009 DE LA COMISIÓN de 19 de octubre de 2009 por el que se ejecuta la Directiva 2007/2/CE del Parlamento Europeo y del Consejo en lo que se refiere a los servicios de red*. OJ L 274, 20.10.2009, pp. 9-18.
- [16] A. Vasiliescu, C. Hauschildt, R. S. Smith, and M. Lutz, "Governance of Persistent Identifiers ,," ARE3NA, D.TD.04, Jul. 2015.
- [17] "Implementation of Identifiers using URIs in INSPIRE - Frequently Asked Questions," *INSPIRE*. [Online]. Available: <http://inspire.ec.europa.eu/print2.cfm//index.cfm/pageid/5120>. [Accessed: 18-Jan-2016].
- [18] Ministerio de la Presidencia, España, *Real Decreto 4/2010, de 8 de enero, por el que se regula el Esquema Nacional de Interoperabilidad en el ámbito de la Administración Electrónica*. 2010.
- [19] España, Ministerio de Hacienda y Administraciones Públicas, *Norma Técnica de Interoperabilidad de Reutilización de recursos de la información*. 2013, pp. 1-27.

## AUTORES

**Francisco J LOPEZ-PELLICER**

*fjlopez@unizar.es*  
Universidad de Zaragoza  
IAAA

**F. Javier ZARAZAGA-SORIA**

*javy@unizar.es*  
Universidad de Zaragoza  
IAAA

**Antonio F. RODRÍGUEZ**

*afrodriguez@fomento.es*  
Centro Nacional de Información  
Geográfica

**Jesús BARRERA**

*jesusb@geoslab.com*  
GeoSLab

**Emilio LÓPEZ**

*elromero@fomento.es*  
Centro Nacional de Información  
Geográfica

**Julián GONZÁLEZ**

*jgonzalezg@fomento.es*  
Centro Nacional de Información  
Geográfica  
**Paloma ABAD**  
*pabad@fomento.es*  
Centro Nacional de Información  
Geográfica