

14C en el sistema IDEArq

WAA

¹⁴C en el sistema IDEArq

Antonio Uriarte González¹,
Carlos Fernández Freire², Alfonso Fraguas Bravo³,
Nuria Castañeda Clemente⁴,
Enrique Capdevila Montes⁵, Ernesto Salas Tovar⁶, Antonio
Gilman⁷, Isabel del Bosque González⁸ y Juan M. Vicent
García⁹

^{1,5,9} Instituto de Historia CCHS - CSIC,

^{2,6,8} Unidad SIG y HD CCHS – CSIC,

⁷California State University – Northridge USA

RESUMEN

IDEArq-C14 es la versión electrónica de la *Base de Datos de Radiocarbono de la Prehistoria Reciente Ibérica*. Se trata de uno de los conjuntos de datos integrados en la *Infraestructura de Datos Espaciales de Investigación Arqueológica* (IDEArq), un sistema de información que da acceso en abierto a conjuntos de datos arqueológicos georreferenciados generados por grupos de investigación del CSIC. Como en toda IDE, su diseño se ajusta a los estándares relativos a la información geográfica y sus datos y servicios están catalogados mediante metadatos normalizados. En este trabajo se presentan aspectos generales de esta plataforma, con especial atención a su estructura conceptual y arquitectura. A partir de ello se explica la forma en la que los datos cronométricos se integran en el conjunto, haciendo énfasis en los aspectos y soluciones funcionales, en particular el modelo de datos, aplicables a cualquier sistema de información cronométrica.

23.1 Introducción

La *Infraestructura de Datos Espaciales de Investigación Arqueológica* (IDEArq)¹ es un sistema de información cuyo objeto es dar acceso a conjuntos de datos arqueológicos generados por grupos de investigación pertenecientes al Consejo Superior de Investigaciones Científicas

¹ IDEArq: <http://www.idearqueologia.org/>. Correo electrónico: idearq@cchs.csic.es

(CSIC)² (Fernández Freire *et alii* 2015; Del Bosque y Vicent 2016; Uriarte *et alii* 2017). La filosofía que articula IDEArq se puede resumir en tres conceptos básicos:

- **Difusión:** Los datos son accesibles en abierto a través de Internet, en la línea de la filosofía *Open Data*³ y cumplen con los criterios FAIR de gestión y administración de datos científicos (Wilkinson *et alii* 2016). Su uso está regulado por una licencia Creative Commons CC BY-NC-SA, que permite utilizar, compartir y adaptar los datos libremente, mencionando la autoría, evitando un uso comercial y manteniendo el mismo tipo de licencia abierta.
- **Armonización:** La información está organizada e integrada mediante un lenguaje común, basado en normas y estándares, como la directiva *Infrastructure for Spatial Information in Europe* (INSPIRE) de la Unión Europea, la serie de normas ISO 19100 y los estándares del *Open Geospatial Consortium* (OGC). Ello garantiza la *interoperabilidad* de datos y servicios, esto es, la posibilidad de que estos sean compartidos con otros sistemas de información.
- **Dimensión espacial:** Los datos están georreferenciados, es decir, están vinculados, directamente a una ubicación geográfica.

En este trabajo se presenta uno de los conjuntos de datos actualmente disponibles en la plataforma IDEArq, la *Base de Datos de Radiocarbono de la Prehistoria Reciente Ibérica* (IDEArq-C14), prestando especial atención al correspondiente modelo de datos y a la arquitectura del sistema en relación con el manejo de los datos cronométricos que contiene.

IDEArq, en su estado actual, ofrece tres conjuntos de datos: el *Corpus de Pintura Rupestre Levantina*⁴ (IDEArq-CPRL), el *Archivo de valores isotópicos de la Península Ibérica $\delta^{13}C$* ⁵ (IDEArq- $\delta^{13}C$) y la *Base de Datos de Radiocarbono de la Prehistoria Reciente Ibérica* (IDEArq-C14), a la que se refiere específicamente el presente trabajo (ver también Uriarte *et alii* 2017).

La *Base de Datos de Radiocarbono de la Prehistoria Reciente Ibérica* consiste en una recopilación crítica de las dataciones radiocarbónicas de la Prehistoria Reciente de la Península Ibérica (España, Portugal y Andorra) y de las informaciones técnicas y contextuales que permiten su utilización como datos cronométricos de referencia. Esta recopilación procede del trabajo individual de Antonio Gilman, llevado a cabo en el marco de sus investigaciones sobre la Prehistoria Reciente de la Península Ibérica (Gilman 1992), un trabajo que se inició en 1982 siguiendo el espíritu del simposio sobre carbono 14 y prehistoria de la Península Ibérica organizado por Martín Almagro Gorbea y Manuel Fernández Miranda (Almagro y Fernández Miranda 1978). Hay que citar también como referencia otras obras de recopilación realizadas en paralelo, entre las que cabe destacar el trabajo de Pedro Castro, Vicente Lull y Rafael Micó

² IDEArq es producto de la colaboración entre el Laboratorio de Arqueología del Paisaje y Teledetección (LabTel) y el Grupo de Investigación Prehistoria Social y Económica (GIPSE), ambos del Instituto de Historia del CSIC (IH), y la Unidad de Sistemas de Información Geográfica y Humanidades Digitales (SIGyHD) del Centro de Ciencias Humanas y Sociales del CSIC (CCHS). El trabajo se inició en el marco del proyecto CONSOLIDER INGENIO 2010 “Investigación en tecnologías para la valoración y conservación del Patrimonio Cultural”, CSD2007-00058. IDEArq fue presentada públicamente el 22 de septiembre de 2016 en el Museo Arqueológico Nacional (Madrid): <https://www.youtube.com/watch?v=2zImFimFR-4&t=1292s>

³ Panton Principles. Principles for Open Data in Science: <http://pantonprinciples.org/>. “By open data in science we mean that it is freely available on the public Internet permitting any user to download, copy, analyse, re-process, pass them to software or use them for any other purpose without financial, legal, or technical barriers other than those inseparable from gaining access to the Internet itself. To this end data related to published science should be explicitly placed in the public domain”.

⁴ Corpus de Pintura Rupestre Levantina (CPRL): <http://www.prehistoria.ceh.csic.es/AAR/>. El CPRL es la versión electrónica del *Corpus de Arte Rupestre Levantino (CARL)* (ver Cruz Berrocal *et alii* 2005). A través de IDEArq es posible acceder a sus fondos fotográficos, así como a información contextual básica, extraída y adaptada del archivo digital original.

⁵ El *Archivo de valores isotópicos de la Península Ibérica ($\delta^{13}C$)* es una base de datos de valores isotópicos de estroncio, oxígeno, carbono y nitrógeno obtenidos de restos humanos y animales recuperados en contextos arqueológicos de la Península Ibérica. La creación de este archivo tiene su origen en el proyecto HAR2013-47776-R, “Dieta y movilidad humana en la Prehistoria de la Península Ibérica (3100-1500 ANE). Los casos de la cuenca media del Tajo y el Alto Guadalquivir”.

sobre cronología de la Prehistoria Reciente de la Península Ibérica y Baleares (Castro *et alii* 1996) y *La base de dades radiocarbòniques de Catalunya*⁶, dirigida por Joan A. Barceló.

El contenido de este repertorio ha sido revisado y normalizado por el equipo editorial de IDEArq, dirigido y supervisado por el propio autor de la recopilación original, con objeto de adaptar sus contenidos al modelo de datos y las demás estructuras informacionales que articulan la plataforma IDEArq. La mayor parte de la información recogida en la base de datos proviene de las fuentes bibliográficas, con la excepción de los datos inéditos procedentes de los archivos del extinto Laboratorio de Geocronología del Instituto de Química Física Rocasolano del CSIC, incorporados gracias a la colaboración de Antonio Rubinos Pérez, responsable de dicho laboratorio hasta su cierre en 2014. Finalmente, desde junio de 2019, está disponible dentro del conjunto de datos la *Base de Datos de dataciones de la Prehistoria Reciente de las Islas Baleares*, elaborada por Rafael Micó Pérez (Universidad Autónoma de Barcelona), Cristina Rihuete Herrada (Museo de Son Fornés) y sus colaboradores. Este repertorio fue parcialmente publicado por Castro, Lull y Micó (1996) y actualizado desde ese momento por los autores citados. En noviembre de 2018, el Dr. Micó transfirió la base de datos al equipo editor de IDEArq-C14 para su integración en el sistema, con objeto de facilitar su acceso a la comunidad científica. Consta de 798 dataciones procedentes de 117 yacimientos, y es el resultado de la aplicación de criterios coherentes con los que han conformado el cuerpo principal del conjunto.

IDEArq-C14 reúne más de 10.000 dataciones radiocarbónicas de más de 2000 yacimientos de la Península Ibérica⁷, número que va aumentando según se incorporan fechas de nuevas publicaciones. El marco cronológico de referencia es la Prehistoria Reciente y la Protohistoria, es decir, las fases cronoculturales comprendidas entre el Epipaleolítico / Mesolítico y la Edad del Hierro. Eventualmente se incorporan dataciones de períodos previos (Paleolítico) o posteriores (épocas romana y medieval), en el caso de los yacimientos que presentan estas fases de ocupación, además de las propias del ámbito cronológico de IDEArq-C14. El potencial de este conjunto de dataciones ya ha sido explotado en varios trabajos sobre la Prehistoria Reciente de la Península Ibérica a escala general (Balsera *et alii* 2015; Lillios *et alii* 2016; Vicent *et alii* e.p.).

Hay que subrayar que, dada la procedencia fundamentalmente bibliográfica de la información, IDEArq-C14 no es un conjunto de datos cronométricos primarios, sino un repertorio de referencias documentales. Su organización en un conjunto estructurado debe permitir a los usuarios finales obtener por sí mismos el acceso a las fuentes de información primaria. Los datos incorporados se refieren únicamente a aquellas informaciones que los responsables de la realización de las dataciones y su publicación original consideraron suficientes para sustentar el valor científico de las mismas. Por ello, el contenido radiocronológico del conjunto de datos es frecuentemente irregular, reflejando las prácticas predominantes en la comunidad arqueológica a lo largo del lento desarrollo de la implantación de las técnicas de referencia en los países ibéricos. Lo que sí es absolutamente homogéneo y obedece a los conjuntos de criterios de normalización de la información científica que han sido citados más arriba, es la presentación de la información documental que contextualiza cada datación y la refiere a sus fuentes originales.

2. IDEArq-C14 en la arquitectura IDEArq

IDEArq es un sistema de información diseñado e implementado como una Infraestructura de Datos Espaciales (IDE). Una Infraestructura de Datos Espaciales (en inglés, *Spatial Data Infrastructure*, *SDI*) es “un sistema informático integrado por un conjunto de recursos (catálogos, servidores, programas, aplicaciones, páginas web,...) que permite el acceso y la gestión de conjuntos de datos y servicios geográficos (descritos a través de sus metadatos),

⁶ La base de dades radiocarbòniques de Catalunya: <http://www.telearchaeology.org/c14/>. Actualmente no operativa, el proyecto ha pasado al entorno GoGet (<http://www.ibererono.org/goget/>) (Nota del Editor).

⁷ A fecha de 6 de agosto de 2019, constaban en la base de datos 10.280 dataciones de 2066 yacimientos.

disponibles en Internet, que cumple una serie normas, estándares y especificaciones que regulan y garantizan la interoperabilidad de la información geográfica”⁸.

La plataforma IDEArq responde a un enfoque abierto, tanto en el acceso a los datos (*Open Data*) como en la arquitectura informática, mediante el uso de software de código abierto (*Open Source*)⁹. Hemos seguido esta misma filosofía en proyectos de IDE anteriores, como ARANO (Fraguas Bravo 2008, 2009) y la IDE del yacimiento de Casa Montero (Madrid)¹⁰ (Fraguas Bravo et alii 2010, 2016). Esta última consiste en la publicación íntegra de los datos primarios generados en la excavación y estudio de materiales.

IDEArq sigue la arquitectura multicapa o multinivel, propia de las infraestructuras de datos espaciales y de otros muchos tipos de sistemas de información. Ello otorga autonomía a cada uno de los componentes del sistema y permite articularlos de forma versátil. Asimismo, reduce el impacto que las modificaciones en un nivel pudieran tener sobre los demás; los sistemas multinivel son fácilmente escalables, es decir, pueden ampliarse y adaptarse a nuevas necesidades sin que ello comprometa su funcionamiento. La arquitectura multicapa más típica, utilizada en IDEArq, consta de tres niveles: datos, servicios y presentación.

3. El nivel de datos

El *nivel de datos* es el relativo al almacenamiento de la información primaria en sí, esto es, a las bases de datos y archivos donde esta está alojada. La información de IDEArq está centralizada en una única base de datos PostgreSQL, de tipo relacional, con la extensión PostGIS para datos geoespaciales. Asimismo, la base de datos, a través de hipervínculos, permite el acceso a documentos digitales archivados en repositorios, como es el caso de las imágenes del *Corpus de Pintura de Rupestre Levantina*.

3.1. El modelo de datos.

La implementación informática de la información del nivel de datos va precedida de la formulación de un modelo conceptual que describe de forma estructurada la realidad que se quiere representar. Un modelo de datos es un conjunto de conceptos, reglas y convenciones que permiten describir una determinada realidad (De Miguel y Piattini 1993: 162).

El universo de discurso de IDEArq es complejo y alberga diferentes niveles, desde los aspectos más genéricos de la información arqueológica hasta los más idiosincráticos de cada conjunto de datos. Por ello, el modelo de datos es escalable, y está organizado en tres niveles básicos:

- Un marco general, establecido *a priori*, que es el relativo al conjunto de datos de *Lugares Protegidos* contemplado en la Directiva INSPIRE, una iniciativa dirigida a la creación de una IDE a escala europea para la armonización y distribución en abierto de la información geográfica creada por los organismos públicos. *Lugares Protegidos* es uno de los conjuntos de datos contemplados en el Anexo I de la Directiva INSPIRE (INSPIRE 2007), que es el relativo a datos de referencia, y se refiere a aquellos lugares que son objeto de conservación en virtud de algún marco legislativo. El modelo conceptual relativo a este conjunto de datos está desarrollado en la *Especificación de Datos de Lugares Protegidos* (INSPIRE Thematic Working Group Protected Sites 2010).
- El *Cultural Heritage Application Schema* (Fernández Freire et alii 2013, 2014), un modelo general sobre entidades culturales que es un desarrollo de la mencionada *Especificación de Datos de Lugares Protegidos* de INSPIRE

⁸ Introducción a las IDE: http://www.idee.es/resources/documentos/Introduccion_IDEE.pdf

⁹ Entre el software de código abierto utilizado, cabe destacar Apache Tomcat como servidor de aplicaciones web, GeoServer como servidor de información geográfica, y PostgreSQL – PostGIS como gestor de bases de datos.

¹⁰ Proyecto Casa Montero: <http://www.casamontero.org/>

y que se elaboró para dar cuenta a nivel abstracto de los lugares protegidos de tipo cultural (Patrimonio Cultural) y de las entidades culturales en general.

- Por último, desarrollos específicos para cada conjunto de datos (por ejemplo, las dataciones radiocarbónicas de la Prehistoria Reciente de la Península Ibérica), con el objeto de dar cuenta de sus aspectos particulares.

Esta estructura modular permite simultáneamente integrar toda la información en un esquema común y, a la vez, tratar los aspectos específicos de cada conjunto de datos tal como ha sido creado. Asimismo, abre la posibilidad de incorporar nuevos conjuntos de datos con características propias.

Aunque no vamos a entrar a describir en detalle la totalidad del modelo conceptual, mostraremos un ejemplo para ilustrar su carácter escalable (Uriarte et alii 2017: 210-211) (Figura 1).

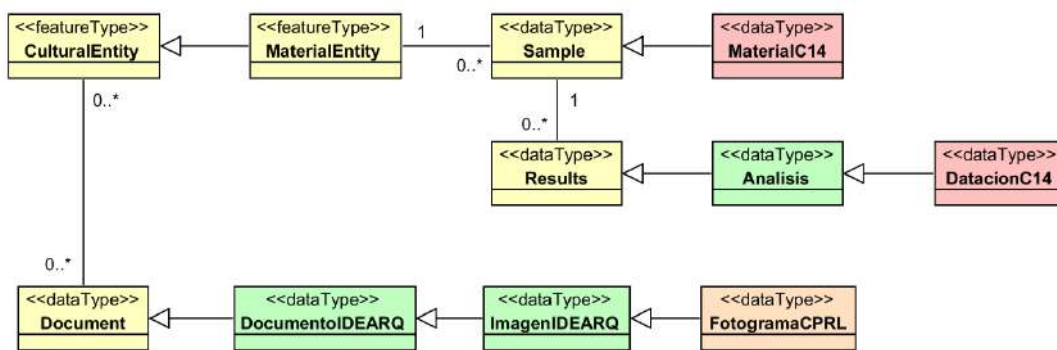


Figura 1: Diagrama de clases UML en el que se muestran algunas de las entidades básicas que forman parte del modelo de datos de IDEArq y las relaciones entre estas (Uriarte et alii 2017).

En su nivel más general o abstracto, el modelo de datos tiene un núcleo conceptual formado por determinadas clases de entidades, relacionadas entre sí. Destacaremos cuatro en lo que se refiere al contenido de este trabajo: *entidad cultural*, *documento*, *muestra* y *análisis*. De ellas, la categoría fundamental es *entidad cultural* y a ella se subordinan las otras. Se define esta como cualquier objeto del mundo real en cuya conformación ha intervenido la acción humana. Dicha definición comprende un amplio abanico de categorías, desde un edificio o un yacimiento arqueológico hasta elementos menores como un muro, un ladrillo, una escultura o un vaso cerámico. Las restantes clases mencionadas son subsidiarias de la de entidad cultural y su naturaleza es eminentemente informacional. Por un lado, un *documento* es cualquier tipo de recurso que contiene información sobre la entidad cultural; por ejemplo, un artículo científico, una fotografía, una planimetría, un vídeo, etc. Por otra parte, la clase *muestra* se refiere a una porción de materia tomada de la entidad cultural y la clase *análisis* a algún tipo de examen científico realizado sobre una muestra y encaminado a obtener algún tipo de datos sobre sus características físico-químicas y/o biológicas. Estas clases más abstractas se complementan, mediante relaciones de herencia, con clases de entidades más concretas, con atributos específicos, destinadas a dar cuenta de los aspectos idiosincráticos de los conjuntos de datos. Por ejemplo, *fotogramas CPRL* es una subclase de documentos dirigida a describir las características específicas de las fotografías que conforman el CPRL. Asimismo, *material C14* y *datación C14* son, respectivamente, subclases de *muestra* y *análisis* y su cometido es integrar en IDEArq la información contenida en la *Base de Datos de Radiocarbono de la Prehistoria Reciente Ibérica*. Recapitulando, este enfoque permite desarrollar el modelo a medida que se incorporan nuevos tipos de datos sin modificar las estructuras preexistentes, de manera que los nuevos modelos de datos serán extensiones del modelo general. Es perfectamente pensable que en el futuro puedan desarrollarse extensiones para acoger tipos de datos que aún no existen y que se incorporarán al registro arqueológico como consecuencia de desarrollos tecnológicos o conceptuales futuros.

Esto significa, tanto a nivel lógico como ontológico, que las dataciones radiocarbónicas no aparecen como entidades en sí, sino que están categorialmente subordinadas a sus correspondientes elementos arqueológicos, conceptualizados como *entidades culturales*. El modelo de datos, por tanto, no tiene como eje las dataciones radiocarbónicas, sino los objetos arqueológicos de los que proceden y de cuyas características físico-químicas dan cuenta, en este caso la *edad isotópica*.

IDEArq-C14, por tanto, aúna esta jerarquía conceptual, integrando la información relativa a las propias dataciones, al contexto arqueológico del que estas proceden y a los documentos en que han sido publicadas. La información es desigual, debido a la heterogeneidad de las fuentes. Por otra parte, está sometida a actualización constante, dada la producción y publicación creciente de nuevas dataciones.

A continuación se describen las entidades que intervienen en la descripción de las dataciones radiocarbónicas, con sus correspondientes atributos (Figura 2). Por razones de claridad expositiva, la descripción está simplificada, omitiéndose todas aquellas entidades y atributos que, aunque existentes en el modelo de datos integral, no aparecen explícitamente en el presente texto. Las entidades implicadas son las siguientes: las dataciones propiamente dichas (*DatacionC14*), los materiales arqueológicos sobre los que se tomaron las dataciones (*MaterialC14*), los yacimientos de procedencia de dichos materiales (incluidos en la entidad genérica *CulturalEntity*) y las referencias bibliográficas relacionadas con cada yacimiento y cada datación (*BiblioIDEARQ*).

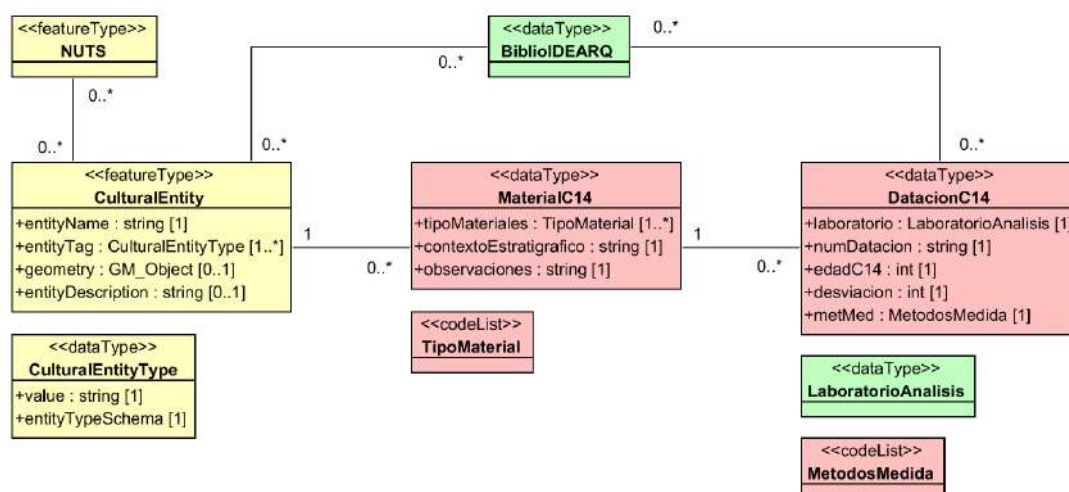


Figura 2: Diagrama de clases UML del modelo de datos de IDEArq-C14 (versión simplificada)

Los yacimientos arqueológicos no están descritos de forma exhaustiva, sino sintética. Se han seleccionado, de la bibliografía de referencia, aquellos datos considerados relevantes para la contextualización e interpretación de los conjuntos de datos asociados. La información considerada es la siguiente:

- Denominación del yacimiento (atributo *entityName*).
- Descriptores tipológicos y crono-culturales (atributo *entityTag*, que toma sus valores de la entidad *CulturalEntityType*). La entidad *CulturalEntityType* reúne todos los descriptores utilizados en IDEArq (atributo *value*), cada uno de los cuales tiene un esquema o listado de pertenencia (atributo *entityTypeSchema*). En este caso, se han utilizado tres esquemas: uno de descriptores cronológicos y dos de descriptores tipológicos, como más abajo se explica.
- Coordenadas geográficas (atributo *geometry*).

- Breve descripción del yacimiento (atributo *entityDescription*), con informaciones básicas de interés y las citas de las referencias bibliográficas relacionadas con el sitio y de las que se ha extraído la información para su descripción.
- Unidades administrativas a las que pertenece el sitio, acorde a la estructura territorial de los estados implicados (España, Portugal y Andorra). En el caso de España y Portugal se ha utilizado la *Nomenclatura de las Unidades Territoriales Estadísticas* (NUTS)¹¹, incorporadas al modelo de datos a través de la entidad *NUTS*. La asignación la realiza automáticamente el propio sistema, en la base de datos geoespacial.

Los descriptores crono-culturales se remiten a las fases generales de la periodización arqueológica convencional de la Península Ibérica (Tabla 1). Cuando no existe consenso sobre la asignación en la literatura científica, se selecciona el valor “Sin atribución cronológica”.

Tabla 1: Descriptores cronológicos de los yacimientos

Período
Paleolítico
Epipaleolítico / Mesolítico
Neolítico
Calcolítico
Edad del Bronce
Edad del Hierro
Romano
Edad Media
Edad Moderna

En cuanto a los descriptores tipológicos, hay dos listados. Uno es cerrado y de carácter general, basado fundamentalmente en criterios funcionales. Consta de tres niveles jerárquicos, ordenados de lo general a lo particular (Tabla 2). Dependiendo de la precisión de la información bibliográfica, se asigna un descriptor de un nivel más general o más específico. La selección de un descriptor de nivel inferior implica la selección automática de los niveles superiores correspondientes. El segundo listado recoge términos tipológicos usados con asiduidad en la bibliografía para referirse a morfologías concretas de yacimientos, normalmente en ámbitos regionales concretos (por ejemplo, *castro*, *talayot* o *motilla*). Dicha lista, al contrario que la tipología general, es abierta y puede admitir nuevos descriptores.

Tabla 2: Descriptores tipológicos de los yacimientos (categorías genéricas)

Nivel 1	Nivel 2	Nivel 3
Hábitat	Edificio aislado	
	Hábitat rupestre	
	Poblado/Asentamiento	Poblado/Asentamiento en llano
		Poblado/Asentamiento en altura

¹¹ NUTS - Nomenclature of territorial units for statistics: <http://ec.europa.eu/eurostat/web/nuts/overview>

Funerario	Enterramiento aislado		
	Necrópolis		
	Megalito funerario		
	Funerario rupestre natural		
	Funerario rupestre artificial		
Uso público	Lugar de representación		
	Termas/Baños		
	Conmemorativo/Monumental		
	Comercial		
	Lúdico		
	Infraestructura		Infraestructura viaria
			Infraestructura hidráulica
			Infraestructura marítima
	Lugar de culto		Edificio de culto
			Lugar de culto rupestre
		Depósito votivo	
Almacenamiento		Silo/Almacén	
Arte rupestre	Cueva		
	Abrigo		
	Aire libre		
Extracción de materias primas	Mina		
	Cantera		
	Salina		
Taller	Taller de cerámica/vidrio		
	Taller metalúrgico		
	Transformación de productos agrícolas/ alimentarios		
	Taller lítico		
	Taller textil		
Defensivo/militar	Campamento		
	Estructura defensiva		
	Campo de batalla		
Otros	Depósito natural		
	Hallazgo aislado		
	Yacimiento negativo/Campo de hoyos		
	Depósito de objetos		
	Megalitismo no funerario		
Sin atribución tipológica			

Los descriptores crono-culturales y tipológicos se combinan en forma de binomios, esto es, asociaciones de un descriptor crono-cultural y otro tipológico. A cada yacimiento se le pueden

asignar cuantos binomios sean necesarios para dar cuenta de todas sus fases cronológicas y de las distintas tipologías asociadas a cada una.

Un aspecto fundamental del modelo de datos es la dimensión geoespacial. Las entidades arqueológicas son entidades en el espacio y como tal son tratadas en IDEArq. En su estado actual, esto se hace a través de los yacimientos arqueológicos. Cada yacimiento ha sido geolocalizado a través de un par de coordenadas que, con mayor o menor precisión dependiendo de cada caso, marcan su ubicación. Se utiliza como sistema de referencia geodésico el ETRS89 (EPSG:4258), que es el oficial a nivel europeo¹². Las coordenadas son de tipo geográfico (longitud y latitud), pudiendo proyectarse a coordenadas UTM en su huso correspondiente.

Las coordenadas de cada yacimiento han sido establecidas a partir de la información disponible sobre él. Su precisión depende de la calidad y accesibilidad de dicha información, por lo que varía de un yacimiento a otro, desde la más alta posible, con la posición ubicada en el área conocida del sitio, hasta grados elevados de imprecisión, en los que la ubicación se remite de manera difusa al paraje.

Del material datado, tenemos los siguientes atributos:

- Material (atributo *tipoMateriales*): Sustancia a partir de la cual se ha obtenido la datación de acuerdo con una lista simplificada. Los valores presentes en esta lista son los siguientes: *Pólenes, Sedimento, Pigmento, Hueso, Madera / Carbón, Fruto / Semilla, Otros macrorrestos vegetales, Concha y Carbonato*. En caso de no disponerse de esta información, se asigna la expresión *Material no referido*.
- Información contextual (atributo *contextoEstratigrafico*): Contexto estratigráfico de la datación.
- Evaluación de la asociación (atributo *evaluacionAsociacion*): Valoración, a efectos cronológicos, de la relación entre el material datado y el contexto estratigráfico que fecha.
- Observaciones (atributo *observaciones*): Aquí se incluye la referencia bibliográfica de la que se ha obtenido la información sobre la datación.

En cuanto a la datación propiamente dicha:

- Laboratorio (atributo *laboratorio*) donde se realizó el análisis. Si no se dispone de este dato, se indica como no referido (*N.R.*).
- Número de serie (atributo *numDatacion*), dentro del laboratorio. Si no se dispone de este dato, se indica como no referido (*N.R.*).
- Fecha media (atributo *edadC14*), en años de C14 antes del presente (BP).
- Desviación típica de la fecha (atributo *desviacion*).
- Método de determinación de la datación (*metMed*): Puede tomar dos valores, que son *Método convencional* y *AMS*. En caso de no disponerse de esta información, se asigna la expresión *Método no referido*.

3.2. Los metadatos

La información en IDEArq está catalogada, esto es, viene descrita por una serie de metadatos. Los metadatos, en su definición más elemental, son datos sobre los datos, esto es, información sobre sus características básicas, como, por ejemplo, su nombre y contenido, su fecha de creación, la institución de la que dependen, cómo han sido elaborados o su extensión espacial y temporal. Los metadatos asociados a un recurso de información permiten al usuario localizarlo, identificarlo, interpretarlo y utilizarlo correctamente.

¹² En el caso español, se adopta el sistema ETRS89 como nuevo sistema de referencia geodésico oficial para la Península Ibérica y las Islas Baleares, sustituyendo al sistema ED50, mediante Real Decreto 1071/2007, de 27 de julio.

La catalogación de la información geográfica está prolijamente regulada por un extenso conjunto de normas y estándares a nivel internacional, europeo y español. El estándar internacional de referencia es la norma ISO 19115 sobre metadatos de información geográfica, publicada en 2003 (TC 211 2003) y recientemente actualizada (TC 211 2014). Dicho estándar ha sido adaptado a nivel europeo en el seno de INSPIRE (Drafting Team Metadata and European Commission Joint Research Centre 2009) y a nivel español a través del *Núcleo Español de Metadatos (NEM)*¹³ (Sánchez Maganto, González Torrado y Juanatey Aguilera 2014).

El uso de estándares propicia la interoperabilidad entre los metadatos de diferentes fuentes de datos geográficas y, por tanto, su integración en catálogos generales.

Los metadatos de los recursos de IDEArq se ajustan al perfil NEM y se han creado con el software CatMDEdit¹⁴. Conforme a los estándares, los metadatos están almacenados en ficheros en formato XML y son accesibles en el propio geoportal IDEArq, además de estar incluidos en el catálogo nacional de la Infraestructura de Datos Espaciales de España¹⁵.

IDEArq ofrece metadatos tanto relativos a los propios datos como a los servicios. Existen metadatos para la capa base de yacimientos y para cada uno de los conjuntos de datos. Asimismo, cada servicio WMS tiene sus propios metadatos.

4. El acceso a la información: nivel de servicios y nivel de usuario

El acceso en abierto a los datos se realiza a través de los niveles de servicios y de usuario. El *nivel de servicios*, mediante protocolos estándar, permite a los usuarios realizar peticiones de información y obtener una respuesta acorde a los parámetros de dicha petición. Los servicios implementados en IDEArq, a día de hoy, son de tipo WMS (*Web Map Service*)¹⁶. WMS es uno de los tipos de servicio definidos por el *Open Geospatial Consortium* (OGC) para la distribución de información geoespacial en la Red. El servicio WMS, en concreto, está concebido para la recuperación de la información en forma de mapas a través de algún interfaz de visualización geográfica habilitado al efecto, bien ligero (visualizador web) bien pesado (software SIG) (Figura 3). Las capacidades (*capabilities*) del WMS permiten al usuario seleccionar de manera interactiva un área de interés y generar un mapa con las entidades contenidas en ella y, opcionalmente, recuperar información temática asociada a estas. Un servicio WMS puede combinarse con otras capas de información geográfica y visualizarse conjuntamente con estas en un mismo mapa. El WMS de IDEArq consta de cuatro capas de información, una general de yacimientos (*idearq*) y una para cada conjunto de datos: *idearq_cpri*, *idearq_c14* e *idearq_dimp*.

¹³ Núcleo Español de Metadatos de Datos (NEM): <http://metadatos.ign.es/web/guest/nem>

¹⁴ CatMDEdit: <http://catmdedit.sourceforge.net/>

¹⁵ Infraestructura de Datos Espaciales de España: <http://www.idee.es/>

¹⁶ Web Map Service: <http://www.opengeospatial.org/standards/wms>

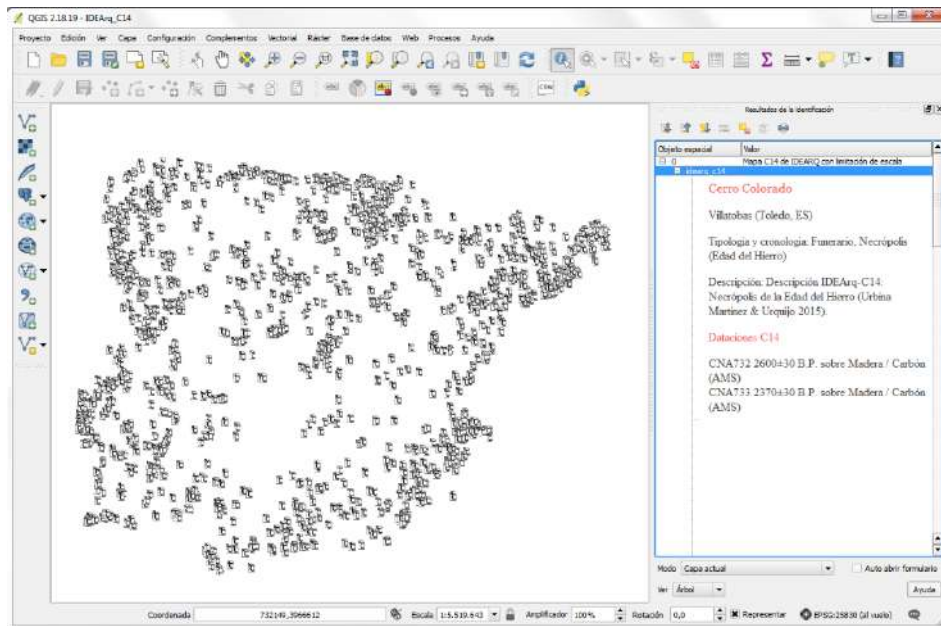


Figura 3: Acceso al WMS de IDEArq-C14 a través de software SIG (QGIS).

El nivel de usuario se refiere a los interfaces a través de los cuales los usuarios pueden acceder al nivel de datos, bien consumiendo los mencionados servicios web (nivel de servicios) bien de forma directa a través de interfaces diseñados *ad hoc*. Los servicios web, como ya se ha dicho, pueden consumirse mediante usuarios ligeros, como visualizadores web preexistentes que disponen de dicha capacidad, o usuarios pesados, como *software* SIG. Por otra parte, existen interfaces creados específicamente para acceder de forma directa a la base de datos geoespacial. El geoportál de IDEArq ofrece dos: el visualizador cartográfico y el interfaz de consulta temática.

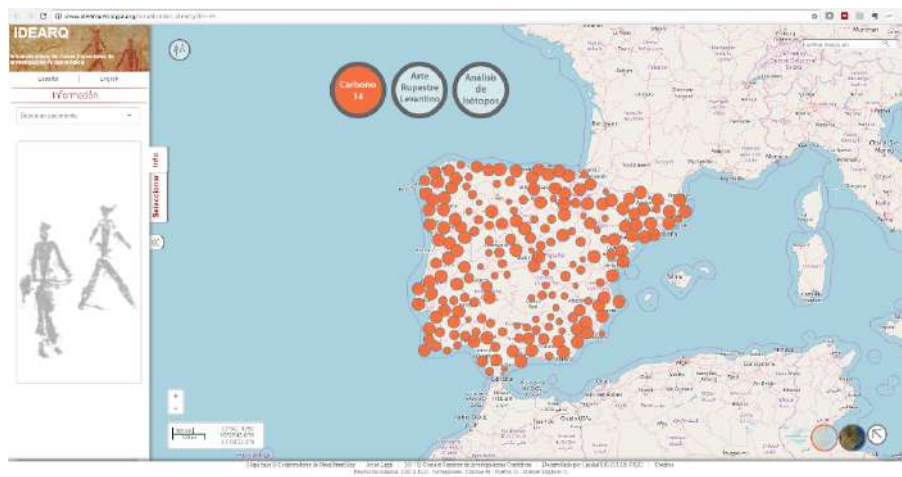


Figura 4: Visualizador cartográfico de IDEArq.

El visualizador cartográfico está basado en *OpenLayers*, una biblioteca *JavaScript* de código abierto para la creación de mapas interactivos en páginas web. Este da acceso a todos los conjuntos de datos alojados en IDEArq, incluida la *Base de Datos de Radiocarbono de la Prehistoria Reciente Ibérica (IDEArq-C14)*. En el visualizador cartográfico, la consulta de las dataciones radiocarbónicas -al igual que el resto de datos temáticos- se hace a través de la capa base de yacimientos (Figura 4), mediante los siguientes pasos:

1. Selección del yacimiento de interés: Por razones de protección del patrimonio arqueológico, IDEARq no muestra las coordenadas de los yacimientos almacenadas en la base de datos, sino que se limita a representar cartográficamente una versión truncada de las coordenadas originales, con un umbral de escala 1:200.000 (Figura 5). En función de la escala de visualización, aquellos yacimientos muy próximos entre sí aparecen agregados mediante un mismo símbolo, cuyo tamaño varía en función del número de yacimientos agrupados. La selección de los yacimientos se puede realizar de dos maneras: 1) seleccionando un yacimiento concreto, bien en el listado desplegable de yacimientos (Figura 6) bien sobre el mapa mediante el cursor; 2) mediante filtros de consulta basados en los descriptores crono-culturales y tipológicos (Figura 7).



Figura 5: Visualizador cartográfico de IDEARq: Zoom máximo a escala 1:200.000.



Figura 6: Visualizador cartográfico de IDEARq: Ventana de búsqueda de yacimientos a través del topónimo



Figura 7: Visualizador cartográfico de IDEARq: Filtros de selección de yacimientos.

2. Consulta de las dataciones del yacimiento seleccionado: Se selecciona el yacimiento de interés para mostrar toda la información asociada a él (Figura 8), entre la que están las dataciones radiocarbónicas (Figura 9). De cada datación se muestran los datos básicos (sigla de laboratorio, la fecha BP con su desviación típica, el tipo de material y el método). Opcionalmente, se pueden visualizar otros dos ítems: el contexto estratigráfico y un apartado observaciones, donde se cita la referencia de la que se ha obtenido la información sobre la datación. Las referencias citadas se muestran a continuación. Asimismo, se ofrece la posibilidad de descargar todas las dataciones del yacimiento con sus atributos básicos.



asociadas a un yacimiento.

El interfaz de consulta temática ha sido específicamente creado para IDEARq-C14 y permite seleccionar subconjuntos de dataciones aplicando determinados criterios. La selección de dataciones se puede realizar de dos maneras:

1. Mediante la selección de un yacimiento concreto en un listado desplegable. Las dataciones resultantes serán todas aquellas procedentes de dicho yacimiento.
2. Mediante la aplicación de filtros basados en determinados atributos, tanto de los yacimientos (regiones geográficas, descriptores crono-culturales y tipológicos) y de los materiales datados (tipo de muestra, tipo de material, taxones biológicos), como de las propias dataciones (fecha media, desviación típica, método de datación, laboratorio).

El resultado de una consulta concreta es el listado de aquellas dataciones que cumplen los criterios seleccionados por el usuario. De cada datación se ofrece toda la información asociada a esta, organizada en datos básicos (fecha media y desviación típica, sigla de laboratorio, tipo de muestra, tipo de material y taxón biológico, método de datación, yacimiento) y complementarios (referencias bibliográficas y resto de atributos sobre el material y la datación). Los datos básicos se pueden imprimir, exportar en formato PDF y copiar y pegar en formato de tabla. El listado de dataciones va acompañado de un visualizador cartográfico donde se muestran los yacimientos de procedencia y un gráfico elaborado a partir de las fechas medias y las desviaciones típicas.

5. Conclusiones

Hemos presentado en este trabajo las características más relevantes del sistema de información cronométrica IDEARq-C14, integrado en la plataforma de información arqueológica IDEARq. El objetivo del proyecto institucional que ha dado lugar a esta plataforma es doble. Por una parte, se trata de facilitar el acceso de la comunidad científica a *corpora* de datos arqueológicos de interés general, tanto pertenecientes a los patrimonios documentales del CSIC, como es el caso del *Corpus de Pintura Rupestre Levantina*, como producidos por la investigación que actualmente se desarrolla en el marco del propio CSIC, como el *Archivo de valores isotópicos de la Península Ibérica*, o por parte de otros investigadores e instituciones, como la *Base de Datos de Radiocarbono de la Prehistoria Reciente Ibérica*, en la que confluyen los resultados de varias fuentes de información.

Pero además de este propósito práctico inmediato, el proyecto IDEARq tiene un objetivo metodológico más general: establecer estándares metodológicos para la articulación práctica de la filosofía *Open Data* en el campo de la Arqueología. En este trabajo hemos presentado los elementos más destacados de la propuesta en lo que se refiere a los planteamientos generales del sistema de información considerado en su conjunto, entre otros su arquitectura y la integración de estándares internacionales y adaptación a las directivas europeas (INSPIRE) para garantizar la interoperabilidad del sistema y accesibilidad de la información sin otras restricciones que las

previstas en el marco jurídico (protección de yacimientos arqueológicos mediante la limitación de escala en la información geoespacial).

En lo que se refiere a la información cronométrica, la aportación metodológica que queremos destacar, al margen de los contenidos de la *Base de Datos de Radiocarbono de la Prehistoria Reciente Ibérica*, que tienen un valor sustantivo para la comunidad científica, es el modelo de datos (MD IDEArq-C14) desarrollado para integrar estos contenidos en el sistema. Consideramos que algunas de las características de éste tienen un valor general orientativo aplicable a cualquier sistema de información cronométrica. Este posible valor general se sustancia en tres puntos fundamentales:

1. Enfatizar la propia relevancia del concepto de “modelo de datos” en la construcción de sistemas de información cronométrica.
2. La concepción de MD IDEArq-C14 como una extensión de un modelo de datos general del registro arqueológico, lo que permite armonizar la lógica interna del sistema de información con el sustrato ontológico al que se refieren los datos.
3. La centralidad de las dimensiones geoespaciales de los datos cronométricos y del marco metodológico definido por el concepto de Infraestructura de Datos Espaciales (IDE) para la construcción de sistemas de información cronométrica.

En lo que se refiere al primer punto, la relevancia de la modelización en el campo de información resulta ser algo no suficientemente explícito en el caso de la mayoría de los sistemas preexistentes. La complejidad del registro arqueológico y las crecientes exigencias planteadas por los criterios de accesibilidad e interoperabilidad en el contexto de la Web 2.0 requieren un esfuerzo de racionalización de las categorías informacionales que se van a usar para representar cualquier elemento de forma que puedan preservarse las relaciones con su contexto que lo hacen científicamente relevante. En el contexto informacional actual no son útiles las bases de datos concebidas como sistemas aislados, en los cuales la complejidad de los datos arqueológicos se trata de representar por la acumulación de atributos. En este sentido, el rasgo más destacado de MD IDEArq-C14 es precisamente que se trata de una extensión de un modelo de datos más amplio, cuyo objeto es el registro arqueológico concebido como una totalidad conceptual. Este hecho permite integrar flexiblemente la información cronométrica en los distintos niveles que definen su contexto arqueológico (objeto datado, contexto arqueológico, yacimiento, clases cronoculturales y/o funcionales de yacimientos).

Finalmente, es crucial la apuesta por el marco conceptual y práctico de las IDE como referencia para la implementación de estos objetivos. Todo dato arqueológico es necesariamente un geodato, y, por consiguiente, la dimensión espacial es el eje que permite la articulación de todos los niveles de generalidad que se han mencionado. Por otra parte, el grado de normalización que se ha alcanzado en el terreno de los sistemas de información geoespacial provee de un marco de referencia que asegura la interoperabilidad de los sistemas y garantiza el acceso a la información que contienen, de acuerdo con los criterios FAIR y las normativas internacionales.

BIBLIOGRAFÍA

- Almagro Gorbea M. y Fernández Miranda M. (eds.) 1978: *C14 y prehistoria de la Península Ibérica*. Fundación Juan March. Madrid.
- Balsera V.; Díaz del Río P.; Gilman A.; Uriarte A. y Vicent J.M. 2015: “Approaching the demography of late prehistoric Iberia through summed calibrated date probability distributions (7000-2000 cal BC)”. *Quaternary Research* 386: 208-211.

- Castro Martínez P.V.; Lull V. y Micó R. 1996: *Cronología de la Prehistoria Reciente de la Península Ibérica y Baleares (c. 2800-900 cal ANE)*. British Archaeological Reports International Series 652. Archaeopress. Oxford.
- Cruz Berrocal M.; Gil-Carles Esteban J.M.; Gil Esteban M. y Martínez Navarrete M.I. 2005: “Martín Almagro Basch, Fernando Gil Carles y el Corpus de Arte Rupestre Levantino”. *Trabajos de Prehistoria* 62 (1): 27-45.
- De Miguel, A. y Piattini, V. 1993: *Concepción y Diseño de Bases de Datos. Del Modelo E/R al Modelo Relacional*. Ra-Ma, Madrid.
- Del Bosque González I. y Vicent García J.M. 2016: “Presentación de IDEArq, Infraestructura de Datos Espaciales de investigación arqueológica”. *Trabajos de Prehistoria* 73 (1): 180-192.
- Drafting Team Metadata and European Commission Joint Research Centre 2009: *INSPIRE Metadata Implementing Rules: Technical Guidelines based on EN ISO 19115 and EN ISO 19119*. European Commission Joint Research Centre.
- Fernández Freire, C.; Del Bosque González, I.; Vicent García, J.M.; Pérez Asensio, E.; Fraguas Bravo, A.; Uriarte González, A.; Fábrega Álvarez, P. y Parcero Oubiña, C. 2013: “A Cultural Heritage Application Schema: Achieving Interoperability of Cultural Heritage Data in INSPIRE”. *International Journal of Spatial Data Infrastructures Research* 8: 74-97.
- Fernández Freire, C.; Parcero Oubiña, C. y Uriarte González, A. (eds.) 2014: *A data model for Cultural Heritage within INSPIRE*. Cadernos de Arqueoloxía e Patrimonio 35. CSIC. Santiago de Compostela.
- Fernández Freire C.; Salas Tovar E.; Gutiérrez González R. y Vicent García J.M. 2015: “A SDI’s in Archaeology in the Iberian Peninsula: Achieving interoperability of Cultural Heritage data in INSPIRE”. *AGILE’2015. International Conference on Geographic Information Science, 9-12 June 2015 - Lisbon, Portugal*.
- Fraguas Bravo, A. 2008: “The ARANO SDI: a spatial data infrastructure for the rock art of Northeast Africa”. *Archaeological Computing Newsletter* 68.
- Fraguas Bravo, A. 2009: *El Arte Rupestre Prehistórico de África Nororiental: Nuevas teorías y metodologías*. CSIC. Madrid.
- Fraguas Bravo, A.; Menchero Fernández, A.; Uriarte González, A.; Vicent García, J.M.; Consuegra Rodríguez, S.; Díaz del Río, P.; Castañeda, N.; Criado, C.; Capdevila, E. y Capote, M. 2010: “Infraestructuras de datos espaciales y datos de excavación arqueológica: SILEX, la IDE de la mina neolítica de sílex de Casa Montero (Madrid)”. *Cuadernos de Prehistoria y Arqueología de la Universidad de Granada* 20: 65-95.
- Fraguas Bravo, A.; Menchero Fernández, A.; Uriarte González, A.; Vicent García, J.M.; Consuegra Rodríguez, S.; Díaz del Río, P.; Castañeda, N.; Criado, C.; Capdevila, E. y Capote, M. 2016: “SILEX. La IDE de la mina neolítica de Casa Montero (Madrid)”. En M^a C. Mínguez García y E. Capdevila Montes (coords.): *Manual de Tecnologías de la Información Geográfica aplicadas a la Arqueología*, Museo Arqueológico Regional de Madrid. Madrid: 361-397.
- Gilman A. 1992: “The Iberian Peninsula, 6000-1500 B.C.”. En R.W. Enrich (ed.) *Chronologies in Old World Archaeology*, University of Chicago Press. Chicago: vol. 1, 295-301; vol. 2, 238-256.
- INSPIRE 2007: “Directive 2007/2/EC of the Council of the European Union and European Parliament Setting up an Infrastructure for Spatial Information in the European Community (INSPIRE). Directive 2007/2/EC”. *Official Journal of the European Union* L108: 1-14.
- INSPIRE Thematic Working Group Protected sites 2010: *INSPIRE Data Specification on Protected Sites – Guidelines v3.1*. European Commission Joint Research Center.
- Lilliois, K.T.; Blanco González, A.; Drake, B.L. y López Sáez, J.A. 2016: “Mid-late Holocene climate, demography, and cultural dynamics in Iberia: A multi-proxy approach”. *Quaternary Science Reviews* 135: 138-153.
- Sánchez Maganto, A.; González Torrado, A. y Juanatey Aguilera, M. 2014: *Núcleo Español de Metadatos (NEM) v. 1.2*. Instituto Geográfico Nacional. Madrid.
- TC 211 2003. *ISO 19115:2003. Geographic information - Metadata*. International Organization for Standardization.
- TC 211 2014. *ISO 19115-1:2014. Geographic information – Metadata – Part 1: Fundamentals*. International Organization for Standardization.

Uriarte González, A.; Fernández Freire, C.; Fraguas Bravo, A.; Castañeda Clemente, N.; Capdevila Montes, E.; Salas Tovar, E.; Gilman, A.; Del Bosque González, I. y Vicent García, J.M. 2017: “IDEArq-C14: Una Infraestructura de Datos Espaciales para la cronología radiocarbónica de la Prehistoria Reciente ibérica”. En J.A. Barceló, J.A.; I. Bogdanovic y B. Morell (eds.): *IberCrono. Cronometrías Para la Historia de la Península Ibérica. Actas del Congreso de Cronometrías Para la Historia de la Península Ibérica (IberCrono 2017). Barcelona, Spain, September 17-19, 2016*, CEUR-WS, Vol-2024 (urn:nbn:de:0074-2024-4): 209-225.

Vicent García, J.M.; Uriarte González, A.; Fernández Freire, C. y Díaz del Río Español, P. e.p.: “Explorando los sesgos en las cronologías radiocarbónicas: La transición Mesolítico-Neolítico en la Península Ibérica a la luz de la *Base de Datos de Radiocarbono de la Prehistoria Reciente Ibérica (IDEArq-C14)* del Prof. Antonio Gilman”.

Wilkinson, M.D. *et alii* 2016: “The FAIR Guiding Principles for scientific data management and stewardship”. *Scientific Data* 3. doi: 10.1038/sdata.2016.18.