

IDEArq-C14: Una Infraestructura de Datos Espaciales para la cronología radiocarbónica de la Prehistoria Reciente ibérica

Antonio Uriarte González (1), Carlos Fernández Freire (2), Alfonso Fraguas Bravo (3), Nuria Castañeda Clemente (4), Enrique Capdevila Montes (5), Ernesto Salas Tovar (6), Antonio Gilman (7), Isabel del Bosque González (8) y Juan M. Vicent García (9)
(1,5,9) Instituto de Historia CCHS - CSIC, (2,6,8) Unidad SIG CCHS – CSIC, (7) California State University – Northridge USA

(1) antonio.uriarte@cchs.csic.es, (2) carlos.ffeire@cchs.csic.es, (3) alfonso.fraguas.bravo@gmail.com, (4) nuriacastanyeda@gmail.com, (5) enrique.capdevila@cchs.csic.es, (6) ernesto.salas@cchs.csic.es, (7) antonio.gilman@csun.edu, (8) isabel.delbosque@cchs.csic.es, (9) juan.vicent@cchs.csic.es

Resumen. IDEArq es un sistema de información concebido para la difusión en Internet de datos arqueológicos a través de su dimensión geoespacial en condiciones de acceso abierto. IDEArq es una Infraestructura de Datos Espaciales (IDE), una plataforma de distribución en abierto de información georreferenciada y diseñada según estándares y normativas internacionales. Los datos pueden consultarse tanto a través de un servicio web consumible desde diferentes aplicaciones (navegadores web, software SIG) como a través del visualizador cartográfico del geoportal de IDEArq. IDEArq-C14 está integrado en IDEArq. Es un archivo de dataciones radiocarbónicas pertenecientes a yacimientos arqueológicos de la Península Ibérica. El ámbito crono-cultural de los datos abarca desde el Epipaleolítico / Mesolítico a la Edad del Hierro, aunque incluye dataciones previas (Paleolítico) o posteriores (épocas romana y medieval) para yacimientos que también presentan estas fases de ocupación. Las dataciones son accesibles a través de los yacimientos a los que están vinculadas. Cada yacimiento está georreferenciado y descrito mediante una serie de atributos básicos (topónimo, unidades administrativas, descriptores crono-culturales y tipológicos, descripción general). A su vez, las dataciones pertenecientes a él se muestran en forma de listado, con los siguientes atributos: sigla de laboratorio, fecha BP con desviación típica, material, método de medida, contexto estratigráfico y observaciones.

IDEArq-C14: A Spatial Data Infrastructure for the Radiocarbon Chronology of Iberian Recent Prehistory

Abstract. IDEArq is an information system designed for the dissemination of archaeological data on the Internet through its geospatial dimension under open access conditions. IDEArq is a Spatial Data Infrastructure (IDE), an open-source georeferenced information distribution platform designed according to international standards and norms. The data can be consulted through a web service from different applications (web browsers, GIS software) and through the cartographic visualizer of the IDEArq geoportal. IDEArq-C14 is integrated in IDEArq. It is a file of radiocarbon dating from archaeological sites in the Iberian Peninsula. The chrono-cultural scope of the data ranges from the Epipaleolithic / Mesolithic to the Iron Age, although it includes previous (Paleolithic) or later (Roman and Medieval) dates for deposits that also present these phases of occupation. The dating is accessible through the sites to which they are linked. Each site is georeferenced and described by a series of basic attributes (place name, administrative units, chrono-cultural and typological descriptors, general description). In turn, the data belonging to it are shown in list form, with the following attributes: laboratory acronym, BP date with standard deviation, material, method of measurement, stratigraphic context and observations.



1 Introducción: Qué es IDEArq

La *Infraestructura de Datos Espaciales de Investigación Arqueológica* (IDEArq)¹ es un sistema de información cuyo objetivo es la difusión a través de Internet de conjuntos de datos arqueológicos georreferenciados producidos por grupos de investigación pertenecientes al Consejo Superior de Investigaciones Científicas (CSIC)² (ver también [Fer15, Bos16]).

IDEArq se vincula al compromiso del CSIC con la política de acceso abierto al conocimiento científico producido con financiación pública, a través de la firma en 2006 de los acuerdos de Berlín de 2003. Asimismo, IDEArq está basada en tecnologías de código abierto.

Una Infraestructura de Datos Espaciales (IDE) se define como “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), disponibles en Internet, que cumple una serie de normas, estándares y especificaciones que regulan y garantizan la interoperabilidad de la información geográfica”³. Por *interoperabilidad* se entiende la capacidad que tienen varios sistemas de compartir información entre sí. La interoperabilidad se consigue mediante la aplicación de estándares y normas, como son, en lo que a los datos georreferenciados respecta, la directiva *Infrastructure for Spatial Information in Europe* (INSPIRE) de la Unión Europea [Ins07], la serie de normas ISO 19100 y los estándares del *Open Geospatial Consortium* (OGC). En esta línea, IDEArq está integrada en la Infraestructura de Datos Espaciales de España (IDEE), gestionada desde el Instituto Geográfico Nacional y englobada a su vez dentro de INSPIRE.

Dada su dimensión geoespacial, en IDEArq se accede a los conjuntos de datos a través de los yacimientos con que se vinculan. En el momento de la redacción de este texto (inicios de 2017), existen dos conjuntos de datos accesibles: la *Base de Datos de Dataciones Radiocarbónicas de la Península Ibérica* y el *Corpus de Pintura Rupestre Levantina* del CSIC (CPRL) [Cru09]⁴. Está previsto incorporar otros conjuntos de datos aportados por diferentes grupos de investigación y laboratorios del CSIC, con el fin de ir construyendo una red de información arqueológica en acceso abierto. Es el caso del *Archivo de valores isotópicos de la Península Ibérica*⁵, 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.

2 Arquitectura general del sistema IDEArq

IDEArq consta de una base de datos en la que se almacena y organiza la información y de dos herramientas de acceso abierto a dicha información por parte de los usuarios a través de la Web: servicio web y visualizador cartográfico.

En la base de datos, la información está estructurada de forma modular, considerando su complejidad y variedad y, a la vez, integrando de forma consistente sus diferentes partes. Los conjuntos de datos integrados en IDEArq son independientes entre sí, pero están estructurados de acuerdo con modelos de datos orientados a objetos que son extensiones de otro general que los hace coherentes y compatibles. Dicho marco conceptual de referencia es el *Modelo de Datos de Patrimonio Histórico* (ver una descripción completa en [Fer14], así como diversas aproximaciones en [Fer12, Fer13] [Par13, Uri13]). Dicho modelo, a su vez, fue elaborado y propuesto como una extensión de la especificación de datos de Lugares Protegidos [Ins10], establecidos como datos de referencia de la Directiva INSPIRE en su Anexo I.

Queda fuera del propósito de este trabajo explicar en toda su extensión el modelo de datos, el cual, no obstante, es uno de los elementos más relevantes del conjunto de la plataforma y es la clave de su funcionalidad. Únicamente señalaremos dos aspectos clave: la conceptualización de la información en una serie de clases básicas con sus respectivos atributos y la escalabilidad del modelo en diferentes niveles de generalidad (Fig. 1). 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

¹ IDEArq: <http://www.idearqueologia.org/>. Correo electrónico: idearq@cchs.csic.es. El contenido de este artículo se refiere a la versión de la plataforma IDEArq presentada en el congreso IberCrono (octubre 2016).

² La plataforma 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 (uSIG) 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 Introducción a las IDE: http://www.idee.es/resources/documentos/Introduccion_IDEE.pdf

⁴ Corpus de Pintura Rupestre Levantina: <http://www.prehistoria.ceh.csic.es/AAR/>

⁵ 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”.



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 Dataciones Radiocarbónicas de la Península 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.

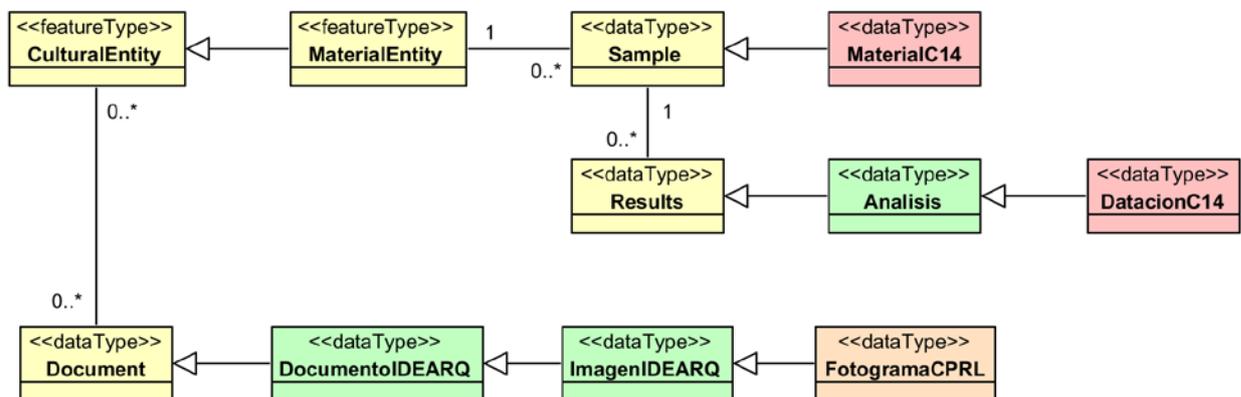


Figura 1 - Diagrama de clases UML en el que se muestran algunas de las entidades básicas que forman parte del modelo de datos de IDEARq y las relaciones entre estas.

Por lo que se refiere a la propia base de datos, es de tipo relacional y está implementada en PostgreSQL, junto con la extensión PostGIS para la gestión de la información geoespacial.

Los dispositivos de acceso abierto a la información son el servicio web y el visualizador cartográfico. El servicio web hace accesible la información a través de diversos interfaces, tanto ligeros (navegadores web) como pesados (software SIG). Es un servicio estándar de tipo WMS (*Web Map Service*), definido por el *Open Geospatial Consortium (OGC)*, y a través de él se puede, mediante mapas interactivos, visualizar los yacimientos y consultar la información asociada a estos. El WMS de IDEARq organiza la información accesible en tres capas: una general de yacimientos (*idearq*), otra que incorpora las dataciones C14 (*idearq_c14*) y una tercera que da acceso a archivo fotográfico del CPRL (*idearq_cpri*).

Tanto los servicios web de IDEARq como los datos contenidos en la base de datos tienen sus propios metadatos, de acuerdo a los estándares de información geográfica: ISO 19115, ISO 19119 y Núcleo Español de Metadatos (NEM). Esto permite incluirlos en los servicios web de catálogo, a través de los cuales es posible conocer las características principales de los mismos, en cuanto a calidad, sistemas de referencia, autoría, linaje, formatos, etc., permitiendo que los datos generados en la investigación científica sean descubribles, accesibles, inteligibles, reutilizables e interoperables con estándares de calidad específicos.

El visualizador cartográfico está integrado en el geoportal de IDEARq (Fig. 2). Está basado en OpenLayers, una biblioteca JavaScript de código abierto para la creación de mapas dinámicos en páginas web. El visualizador da acceso a los datos temáticos a través de la capa base de yacimientos. Los yacimientos se pueden consultar de dos maneras: 1) seleccionando un yacimiento concreto, bien en un listado de yacimientos bien sobre el mapa mediante el cursor; 2) mediante filtros de consulta basados en descriptores crono-culturales y tipológicos (Fig. 3).

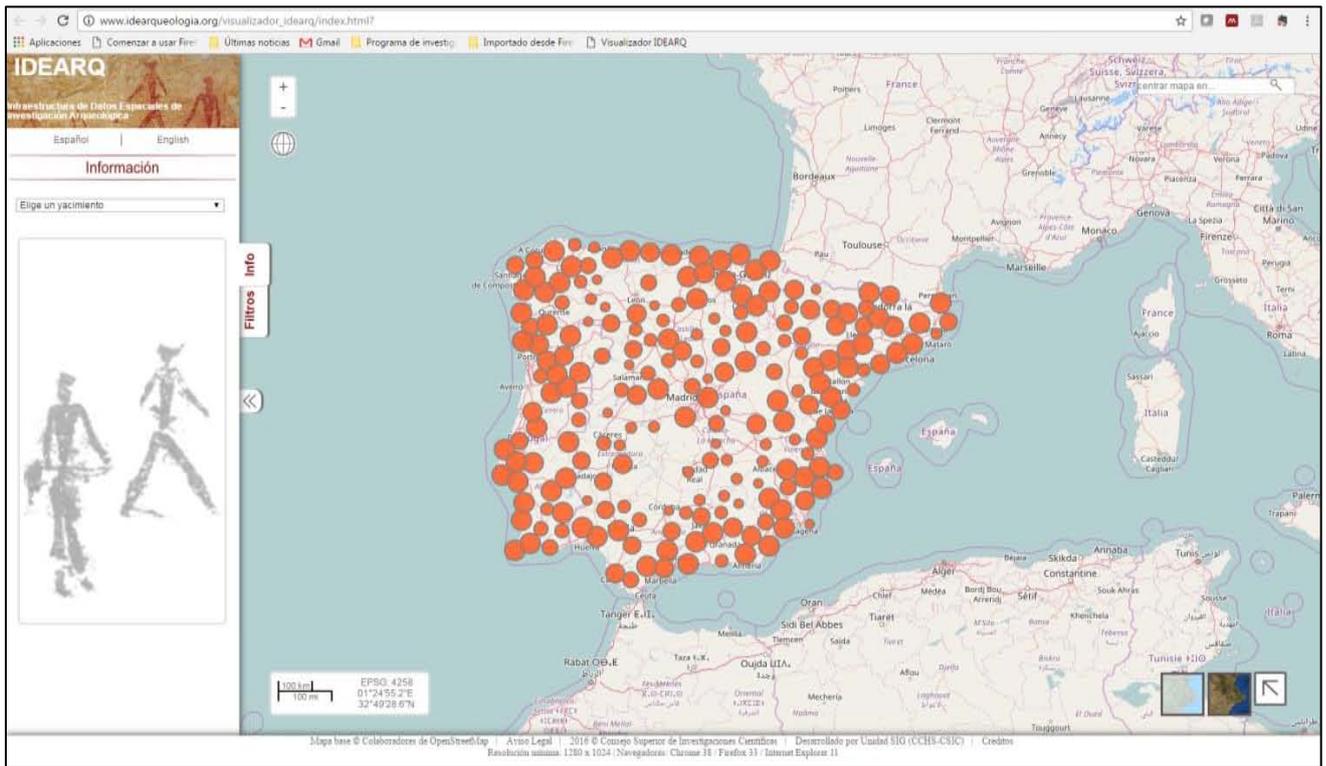


Figura 2 - Visualizador cartográfico de IDEARq.

Filtros

Conjuntos de datos Todos | Ninguno

Carbono 14 <small>(bibliografía)</small>	Arte Rupestre <small>(bibliografía)</small>
---	--

Cronologías Todos | Ninguno

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

Tipologías Todos | Ninguno

Defensivo/Militar	Transformación de materias primas
Extracción de materias primas	Uso público
Estructuras públicas	Culto
Funerario	Hábitat
Almacenamiento	Arte Rupestre
Lugar de función indeterminada	

Filtros Info

Figura 3 - Visualizador cartográfico de IDEARq: Filtros de selección de yacimientos.



IDEArq presenta los datos fundamentales contenidos en la base de datos, así como las referencias necesarias para que los usuarios puedan localizar las fuentes primarias de información, a las que pueden acudir para contextualizarlos. En la mayor parte de los casos, sin embargo, la base de datos contiene información adicional que no se muestra en la plataforma, y a la que se puede (o podrá acceder) a través de interfaces de consulta temática específicos, ya en servicio (como en el caso del CPRL) o que lo estarán en el futuro.

3 La capa base de yacimientos

Como se ha dicho arriba, el acceso a los conjuntos de datos se hace a través de los yacimientos arqueológicos a los que estos se vinculan. Los yacimientos arqueológicos se presentan como una capa georreferenciada. Esta capa solo contiene sitios arqueológicos registrados en alguna de las bases de datos, y no puede considerarse como una representación exhaustiva del registro arqueológico de la Península Ibérica. Por tanto, la capa de yacimientos está formada por aquellos sitios arqueológicos para los que existen datos en alguna de las capas temáticas integradas en la plataforma. En el momento de elaboración de este texto, tiene 1896 yacimientos, 1803 de ellos con información sobre dataciones C14 (Fig. 4).

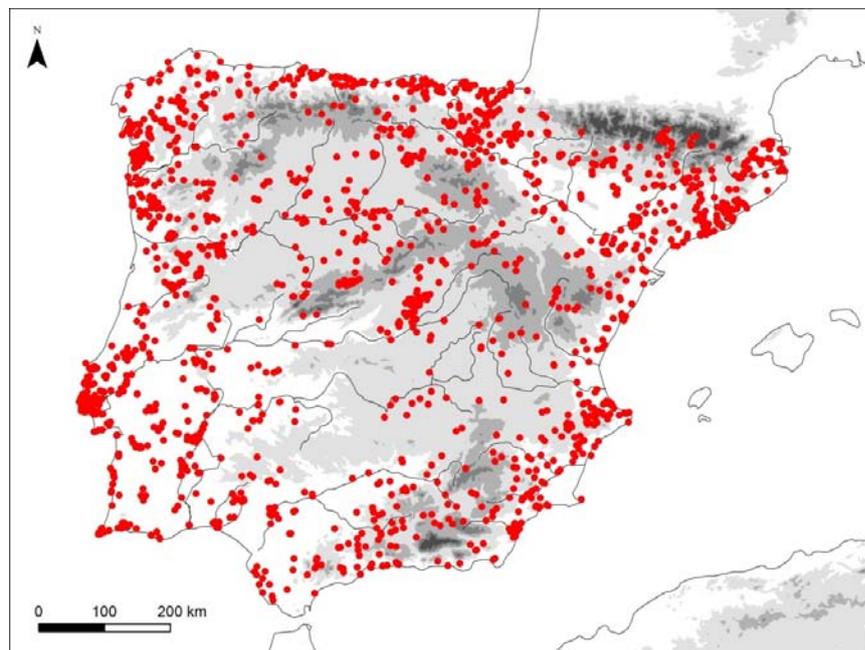


Figura 4 - Yacimientos con dataciones radiocarbónicas incluidos en IDEArq.

3.1 Información geoespacial: Localización

La georreferenciación de los yacimientos se ha basado en el sistema de coordenadas elipsoidales ETRS89 (EPSG:4258); todas las coordenadas obtenidas en las fuentes de información manejadas han sido transformadas a este sistema. El Real Decreto 1071/2007 establece ETRS89 como sistema geodésico oficial en España para la referenciación geográfica y cartográfica en el ámbito de la Península Ibérica y las Islas Baleares. Por esta razón es el utilizado por la propia IDE de España.

Para la localización de los yacimientos se han utilizado las siguientes fuentes de información:

- Registros oficiales, por ejemplo IDE y bases de datos institucionales o cartas arqueológicas de los servicios de patrimonio de las diferentes entidades administrativas.
- Referencias geográficas precisas presentadas en la bibliografía de referencia, bien como coordenadas numéricas, bien en forma de mapa.
- Geolocalización sobre las imágenes del Plan Nacional de Ortofotografía Aérea (PNOA), generalmente a través del visualizador Iberpix, a partir de informaciones imprecisas (descripciones literarias y/o croquis) presentadas por la bibliografía de referencia.
- Localización directa: búsqueda del emplazamiento más probable a partir de la información disponible en la bibliografía sobre toponimia, accidentes geográficos, etc.

El método de localización ha consistido en el cotejo de estas fuentes con objeto de conseguir la aproximación más precisa posible. Siempre que ha sido posible se han combinado todas las fuentes disponibles, valorando comparativamente



los datos que proporcionan. La precisión de las coordenadas varía según los yacimientos en función de la información disponible para cada uno. La localización se ha llevado a cabo a través de los siguientes pasos:

- Selección de la mejor fuente posible entre las disponibles, es decir, aquella que proporciona una información más precisa.
- Búsqueda directa del yacimiento sobre las fuentes cartográficas (ortofotografías aéreas) a partir de los datos procedentes de la fuente.
- Revisión del emplazamiento seleccionado cotejando el resto de las fuentes.

Aunque las coordenadas están almacenadas en la base de datos con sus valores exactos, su presentación en la Web, tanto a través del servicio WMS como del visualizador cartográfico, está sujeta a límites por motivos de seguridad. Se ha establecido un umbral de aproximación en la escala 1:200.000, más allá del cual no se pueden visualizar los yacimientos (Figura 5). Además, todas las coordenadas están redondeadas a 1 km tanto en el WMS como en el visualizador.

Para terminar, hay que señalar que la calidad de la información espacial depende en gran parte de la de las fuentes disponibles. Así por ejemplo, de los 1789 yacimientos para los cuales tenemos (en el momento de redactar estas páginas) al menos una fecha C14 con una edad BP, hay 233 (un 13%) para los que no ha sido posible determinar una localización con precisión de 100 metros. Probablemente, esta proporción se mantendrá a medida que la muestra aumente, y representa una característica de la información arqueológica ibérica globalmente considerada. En cualquier caso, el mencionado umbral de escala impuesto a la información espacial por razones de seguridad, limita el alcance práctico de estas limitaciones para los usuarios.

3.1 Descripción

Cada sitio arqueológico aparece descrito por una serie de atributos. Dicha descripción se ha elaborado a partir de la información contenida en las fuentes existentes sobre el yacimiento, fundamentalmente las referencias bibliográficas. Los editores de IDEArq se han limitado a sintetizar dicha información, prestando especial atención a aquellos datos considerados relevantes para la interpretación de los conjuntos de datos asociados. Generalmente se ha recurrido a la referencia que se ha considerado más representativa, preferentemente la que recoge la versión de los excavadores y/o los investigadores que han estudiado y publicado los materiales. Normalmente se ha dado prioridad a las referencias más recientes y/o más detalladas sobre las más limitadas o parciales. Por ejemplo, se han preferido las memorias de excavación a las noticias preliminares o discusiones generales en las que se alude a dicho sitio.

La información ofrecida sobre cada yacimiento es la siguiente:

- Denominación del yacimiento, de acuerdo con la bibliografía de referencia.
- Términos relativos a las unidades administrativas a las que se asigna el sitio.
- Descriptores tipológicos y crono-culturales de las diferentes fases de uso registradas.
- Breve descripción del yacimiento, con informaciones básicas de interés y citas bibliográficas de referencia. La bibliografía citada puede encontrarse en listas bibliográficas accesibles desde el propio visualizador.

Las unidades administrativas, jerárquicamente ordenadas, obedecen a la estructura territorial de España, Portugal y Andorra. La asignación de cada yacimiento a la correspondiente cadena de unidades administrativas se realiza en la propia base de datos geoespacial, asignando a cada sitio la unidad administrativa que lo contiene. En el caso de España y Portugal se ha recurrido a la *Nomenclatura de las Unidades Territoriales Estadísticas (NUTS)*⁶.

Los descriptores crono-culturales se remiten a las fases generales de la periodización arqueológica convencional de la Península Ibérica (Tabla 1), omitiendo los términos relativos a subfases (por ejemplo “Bronce Medio”) o a periodizaciones regionales (por ejemplo “Cultura de El Argar” o “argárico”). Se ha simplificado al máximo la lista de descriptores unificando bajo una sola denominación términos equivalentes que han sido usados en la bibliografía. Por ejemplo, se ha utilizado “Calcolítico” en lugar de “Eneolítico” o “Edad del Cobre”, aunque estos términos aparezcan en la bibliografía de referencia. Asimismo, se ha considerado preferible utilizar el descriptor “Epipaleolítico / Mesolítico” para englobar los yacimientos precerámicos del Holoceno sin prejuzgar interpretaciones que, sin embargo, pueden recuperarse fácilmente en la bibliografía de referencia. En los casos en los que la asignación crono-cultural de un yacimiento sea ampliamente debatida y no exista consenso entre los especialistas, se ha optado por la expresión “Sin atribución cronológica”. En todo caso, las asignaciones se basan en los contenidos de las referencias bibliográficas, escogiéndose los términos más próximos a los utilizados por los autores.

Un segundo grupo de descriptores son los tipológicos, que pueden variar a lo largo de las sucesivas fases de ocupación. Hay dos listados de descriptores tipológicos. Uno es de carácter general y cerrado, basado fundamentalmente en criterios funcionales. Consta de tres niveles, ordenados jerárquicamente de lo más general a lo más particular (Tabla 2). Cada categoría general (nivel 1) puede incluir varias categorías de nivel intermedio (nivel 2) que, a su vez, pueden desdoblarse en categorías específicas (nivel 3) correspondientes a funcionalidades concretas dentro de cada tipo. La selección de un descriptor de nivel inferior implica la selección de los niveles superiores correspondientes. Sólo se asignan descriptores de nivel inferior cuando hay elementos de juicio en la bibliografía para precisar la asignación funcional.

⁶ NUTS - Nomenclature of territorial units for statistics: <http://ec.europa.eu/eurostat/web/nuts/overview>



Tabla 1: Descriptores cronológicos de los yacimientos

Periodo
Paleolítico
Epipaleolítico / Mesolítico
Neolítico
Calcolítico
Edad del Bronce
Edad del Hierro
Romano
Edad Media
Edad Moderna
Sin atribución cronológica

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

Nivel 1	Nivel 2	Nivel 3		
Hábitat	Edificio aislado	Poblado/Asentamiento en llano Poblado/Asentamiento en altura		
	Hábitat rupestre			
	Poblado/Asentamiento			
Funerario	Enterramiento aislado			
	Necrópolis			
	Megalito funerario			
	Funerario rupestre natural			
	Funerario rupestre artificial			
Uso público	Lugar de representación	Infraestructura viaria Infraestructura hidráulica Infraestructura marítima Edificio de culto Lugar de culto rupestre Depósito votivo Silo/Almacén		
	Termas/Baños			
	Conmemorativo/Monumental			
	Comercial			
	Lúdico			
	Infraestructura			
	Lugar de culto			
	Almacenamiento			
	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		

El segundo listado recoge términos tipológicos usados ampliamente en la bibliografía para describir morfologías concretas de yacimientos, normalmente en ámbitos regionales restringidos (por ejemplo, *castro*, *talayot* o *motilla*). Dicha lista, al contrario que la tipología general, es abierta y puede admitir nuevos descriptores si así lo propicia la incorporación de nueva información a IDEArq. Estos términos complementan a la tipología funcional jerárquica, pero no la sustituyen.

Los descriptores crono-culturales y tipológicos se muestran 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.

Además, en el visualizador cartográfico los descriptores se utilizan como filtros para la selección de yacimientos, concretamente los crono-culturales y los del nivel 1 del listado tipológico general (Fig. 3). Estos filtros se complementan con un tercero, que filtra los yacimientos según a qué conjunto o conjuntos de datos está vinculado.

4 C14 en IDEArq: Base de Datos de Dataciones Radiocarbónicas de la Península Ibérica (IDEArq-C14)

La *Base de Datos de Dataciones Radiocarbónicas de la Península Ibérica* se basa en la recopilación crítica iniciada en 1982 por Antonio Gilman como parte de sus investigaciones sobre la Prehistoria reciente de la Península Ibérica [Gil92] y al hilo del simposio sobre C-14 y prehistoria de la Península Ibérica organizado por Martín Almagro Gorbea y Manuel Fernández-Miranda [Alm78].

La mayor parte de la información recogida en la base de datos proviene de las fuentes bibliográficas, por lo que refleja el estado de la cronología radiocarbónica de la Península Ibérica aceptado por la comunidad científica. A ella se unen datos inéditos procedentes de los archivos del extinto Laboratorio de Geocronología del Instituto de Química Física Rocasolano del CSIC, obtenidos gracias a la generosidad de Antonio Rubinos Pérez, responsable de dicho Laboratorio hasta su cierre en 2014. Asimismo, cabe señalar la mucha ayuda prestada por obras de recopilación realizadas en paralelo, entre las que destacan el trabajo de Pedro Castro, Vicente Lull y Rafael Micó sobre Cronología de la Prehistoria Reciente de la Península Ibérica y Baleares [Cas96] y *La base de dades radiocarbòniques de Catalunya*⁷, organizada por Joan A. Barceló.

La labor de recopilación ha combinado la localización de los recursos bibliográficos, la extracción de la información de acuerdo con criterios normalizados, la lectura crítica de la información ofrecida en las publicaciones, la resolución de contradicciones y ambigüedades, y el cotejo de las diferentes fuentes existentes. Hay que subrayar que, por la propia naturaleza de las fuentes de información manejadas, el contenido de la base de datos es forzosamente irregular, como lo son los niveles de detalle y calidad de la información bibliográfica utilizada. En este sentido, las lagunas de información son más abundantes en las dataciones procedentes de las referencias más antiguas. No se han incluido en la base de datos aquellas referencias en las que resulta imposible resolver contradicciones o rellenar omisiones que inutilizan su uso como dato científico contrastado.

IDEArq-C14 recoge, en el momento de la redacción de este trabajo, 8450 fechas C14 de un total de 1803 yacimientos. En su estado actual, no es un repertorio exhaustivo de la cronología radiocarbónica de la Península Ibérica. Se restringe a la Prehistoria Reciente y la Protohistoria, es decir, a los periodos comprendidos entre el Epipaleolítico / Mesolítico y la Edad del Hierro. Eventualmente incorpora dataciones previas (Paleolítico) o posteriores (épocas romana y medieval) en el caso de yacimientos que presentan estas fases de ocupación además de las incluidas en el ámbito propio de la base de

⁷ La base de dades radiocarbòniques de Catalunya: <http://www.telearcheology.org/c14/>



datos. Se ha procedido así para proporcionar a los usuarios la totalidad de los contextos cronológicos de estos yacimientos. La mayoría de las dataciones incluidas se distribuyen en el intervalo de fechas convencionales entre el 10.000 y el 1.500 BP. Igualmente, el ámbito geográfico está restringido al territorio de la Península Ibérica, aunque está previsto incluir en el futuro los territorios insulares españoles (Baleares y Canarias). Este conjunto de dataciones ha sido usado en dos trabajos sobre la Prehistoria Reciente de la Península Ibérica [Bal15, Lil16], mostrándose como un ejemplo de la utilidad de este tipo de repertorios en los estudios arqueológicos a escala general.

A través del servicio WMS y el visualizador cartográfico se puede acceder a la mayor parte de la información almacenada en la base de datos. Como se ha mencionado más arriba, no todos los atributos son accesibles. El objetivo de la presentación de las dataciones en el visualizador es permitir a los usuarios localizarlas y ofrecerles las referencias necesarias para ampliar la información acerca de las mismas en sus fuentes originales. Por ello, se han seleccionado, para ser accesibles a través del visualizador y los servicios, los datos básicos sin los cuales una datación C14 no puede considerarse un dato científico útil en un contexto de aplicación estándar.

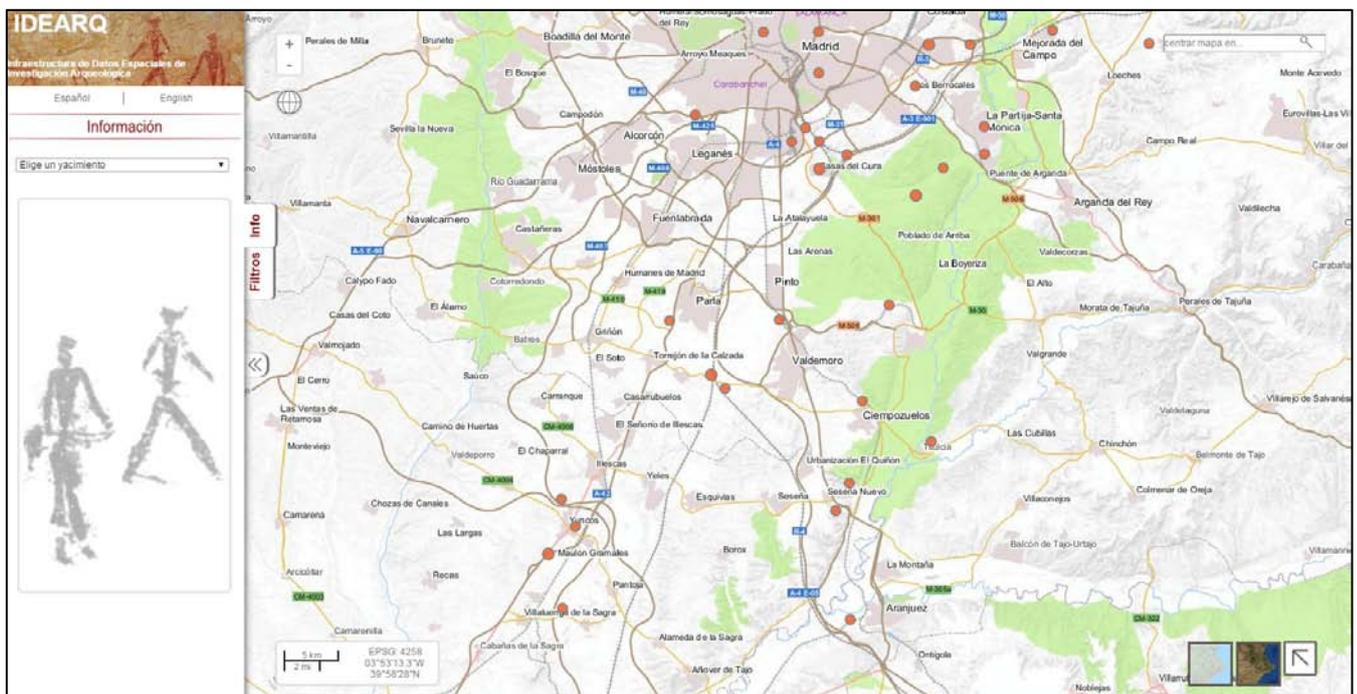


Figura 5 - Visualizador cartográfico de IDEARq: Zoom máximo a escala 1:200.000.

Cada datación se muestra con los siguientes atributos (Fig. 6):

- Laboratorio y número de serie: Si no se dispone de alguno de estos datos, se indica como no referido (N.R.).
- Datación: Se anota siempre en años de C14 antes del presente (BP) con su correspondiente desviación típica. En algunos casos no se dispone de esta información; no obstante, se ha incluido la datación porque sí se dispone de información de interés en otros atributos.
- Material: 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*.
- Método de determinación de la datación: 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*.
- Información contextual: Contexto estratigráfico de la datación y observaciones (normalmente la referencia bibliográfica de la datación).

Asimismo, es posible descargar la información básica de las dataciones en formato csv, con los siguientes atributos: laboratorio y número de serie, valor de la datación BP con su desviación típica, tipo de material y método de medida.

5 IDEARq-C14 y el registro cronométrico de la Prehistoria Ibérica

Posiblemente, el uso más frecuente de un sistema de información como IDEARq-C14 sea la consulta de datos concretos o series de datos. El sistema permite al usuario obtener la información necesaria para contextualizar dataciones individuales



o series de dataciones de acuerdo con criterios selectivos espaciales, crono-culturales o tipológico-funcionales. Cabe preguntarse, no obstante, en qué medida el conjunto de datos recoge íntegramente la masa total de las dataciones realizadas y, por lo tanto, es representativo del registro radiocarbónico peninsular.

El problema es que no existe un término de comparación objetivo e independiente que permita evaluar esta cuestión. El conjunto de datos IDEArq-C14 es, hoy por hoy, la más extensa recopilación de dataciones radiocarbónicas de la prehistoria peninsular. Por ello no es posible resolver la cuestión de su representatividad en términos objetivos, aunque sí lo es establecer, si es el caso, la existencia de sesgos o limitaciones que deban ser tenidos en cuenta a la hora de utilizarlo. Este es el objetivo de estos comentarios finales.

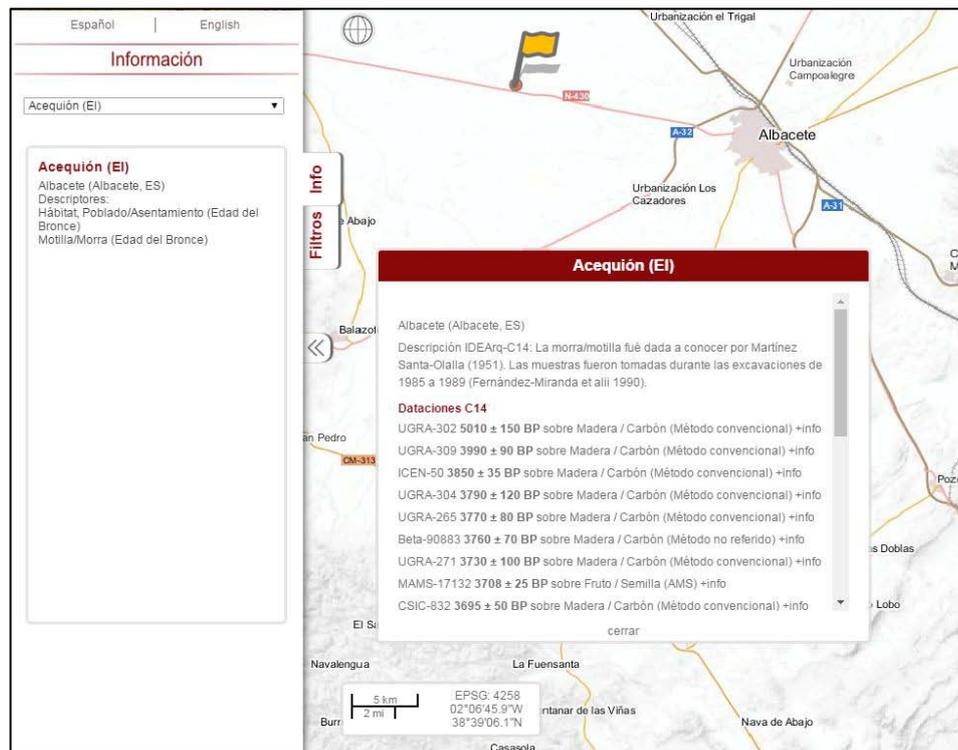


Figura 6 - Visualizador cartográfico de IDEArq: Información sobre un yacimiento y las dataciones radiocarbónicas asociadas a este.

Como se ha dicho, el trabajo sistemático de Gilman, sostenido durante 35 años, constituye el núcleo de los contenidos de IDEArq-C14, cuyo equipo, dirigido por él mismo, se ha limitado a adaptar la presentación de los datos a las disposiciones del modelo de datos, así como a completar y normalizar los sistemas de referencias y descriptores, revisar la geolocalización y realizar otras tareas de edición.

El método de recopilación de la información establece por sí mismo una primera limitación que debe ser tenida en cuenta a la hora de valorarla: IDEArq-C14 contiene casi exclusivamente datos procedentes de publicaciones accesibles para un investigador a través de su práctica profesional. Es decir, que representa el conjunto de la información radiocarbónica publicada en revistas científicas o monografías accesibles a través de las instituciones normales que estructuran la práctica arqueológica, incluyendo entre estas, junto a las bibliotecas especializadas y universitarias, las redes de intercambio profesional normales, tanto personales como académicas. Desde este punto de vista, podemos suponer que publicaciones de escasa difusión (como boletines locales, por ejemplo) han podido quedar fuera del foco. En este sentido, puede decirse que IDEArq-C14 representa lo que Thomas Kuhn [Kuh62] llamaría “Ciencia normal” en lo que se refiere a la cronología radiocarbónica peninsular, puesto que, a efectos del estado de la misma, esas publicaciones difícilmente accesibles no desempeñan ningún papel.

Junto a las limitaciones derivadas de la accesibilidad de las fuentes, cabría esperar que los intereses y campos de especialización del responsable de la recopilación limitasen en algún sentido la representatividad del conjunto de datos. En efecto, los intereses profesionales de Gilman se centran en el cuadrante sureste de la Península del Neolítico al Bronce.

Este hecho se refleja, en primer término, en la explícita acotación crono-cultural del ámbito de representatividad del conjunto de datos al lapso Epipaleolítico / Mesolítico - Edad del Hierro. Como se ha dicho más arriba, se han incluido dataciones anteriores o posteriores a estos límites solo en el caso de series en yacimientos multiestratificados, con objeto de proporcionar al usuario la totalidad del contexto arqueológico. Pero, en última instancia, el conjunto de datos IDEArq-C14 se refiere exclusivamente a la Prehistoria Reciente ibérica en condiciones representativas, es decir, no sesgadas. Los



datos que caen fuera de este ámbito están, por lo tanto, necesariamente sometidos a condiciones selectivas: proceder de yacimientos que presentan alguna fase de ocupación entre el Epipaleolítico / Mesolítico y la Edad del Hierro. Como datos aislados pueden ser útiles, pero su conjunto no representa insesgadamente el registro radiocarbónico general de los periodos correspondientes (Paleolítico y periodos históricos).

No obstante, dentro del intervalo Epipaleolítico / Mesolítico - Edad del Hierro pueden presentarse otro tipo de sesgos, que podemos llamar “locales”, debidos a la dificultad de extender el mismo grado de conocimiento a las fuentes y, en general, a la práctica arqueológica de todos los contextos regionales y cronológicos de la Península. Por ejemplo, es posible que pudiera haber deficiencias con respecto a la Edad del Hierro del noroeste, por citar el caso más alejado posible del núcleo de los intereses profesionales del compilador. Estas deficiencias, caso de existir, sólo podrían ser establecidas por especialistas en los registros regionales de los distintos periodos, y cabe esperar que la publicación del conjunto de datos permita su corrección mediante el concurso cooperativo de la comunidad científica.

Ahora bien, la cuestión primordial es establecer si estos “sesgos locales” menoscaban la representatividad del conjunto de datos globalmente considerado o, por el contrario, puede mantenerse la hipótesis de que IDEArq-C14 es una muestra no sesgada del registro radiocarbónico peninsular dentro de los límites crono-culturales mencionados.

Creemos que esto último puede mantenerse razonablemente por varias razones. En primer lugar, porque la información contenida en el conjunto de datos es coherente con la historia de la formación del propio registro, y permite reconstruir con bastante detalle aspectos concretos de la misma. En segundo lugar, porque la variabilidad interna del conjunto de datos es coherente con hipótesis arqueológicas independientes de su contenido sustantivo. Examinaremos brevemente estas dos cuestiones, aunque, naturalmente, requieren un tratamiento más profundo de lo que podemos permitirnos aquí.

En lo que se refiere a la historiografía de la aplicación del C14 en el ámbito de la Prehistoria peninsular, sirva como ejemplo la comparación de los números de yacimientos con dataciones en 1978 [Alm78], 1996 [Cas96] y en IDEArq (excluyendo los yacimientos baleáricos y los exclusivamente paleolíticos en los listados anteriores). La progresión es aritmética: a intervalos de unos 20 años el número se multiplica aproximadamente por 4 (Fig. 7). Y la pauta es parecida con respecto al número de fechas, excepto que el factor de aumento es superior a 5 (Fig. 8). Es interesante también observar cómo se va rellenando la Península de yacimientos fechados. En 1978 la dispersión refleja, por una parte, el interés en la secuencia clásica del sureste y, por otra, la energía de algunos grupos de investigación (Fig. 9). Es decir, no puede considerarse representativa de la distribución de yacimientos en la Prehistoria. En 1996 el mapa empieza a rellenarse, pero el impacto de ciertos grupos de investigación sigue siendo evidente (véase el distrito de Porto, por ejemplo) (Fig. 10). La situación presente todavía no puede considerarse representativa de la distribución real de yacimientos pre- y protohistóricos (Fig. 11). Se nota, por ejemplo, el contraste entre las principales ciudades de la Península comparadas con demarcaciones administrativas vecinas. También se ve reflejada la importancia historiográfica del sureste y el impacto de ciertos grupos de trabajo.

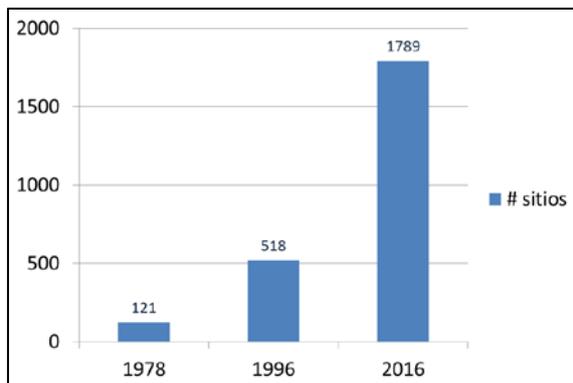


Figura 7 - Evolución del número de yacimientos con dataciones C14

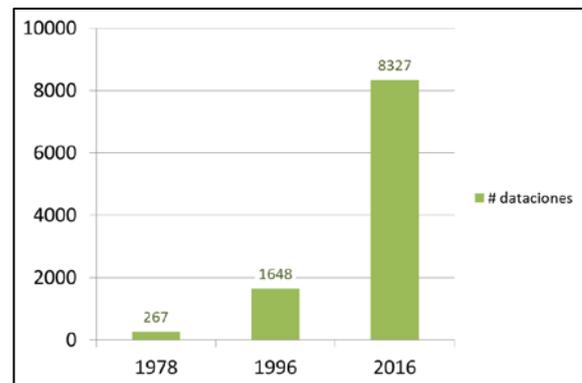


Figura 8 - Evolución del número de dataciones C14



En resumen, los datos de IDEArq-C14 no solo permiten reconstruir en líneas generales la progresión de la aplicación del C14 en la arqueología prehistórica ibérica, sino que ponen de manifiesto algunos hechos relevantes de la historia de la investigación, como son la actividad de grupos concretos o la incidencia de los fenómenos de urbanización sobre la formación del registro arqueológico.

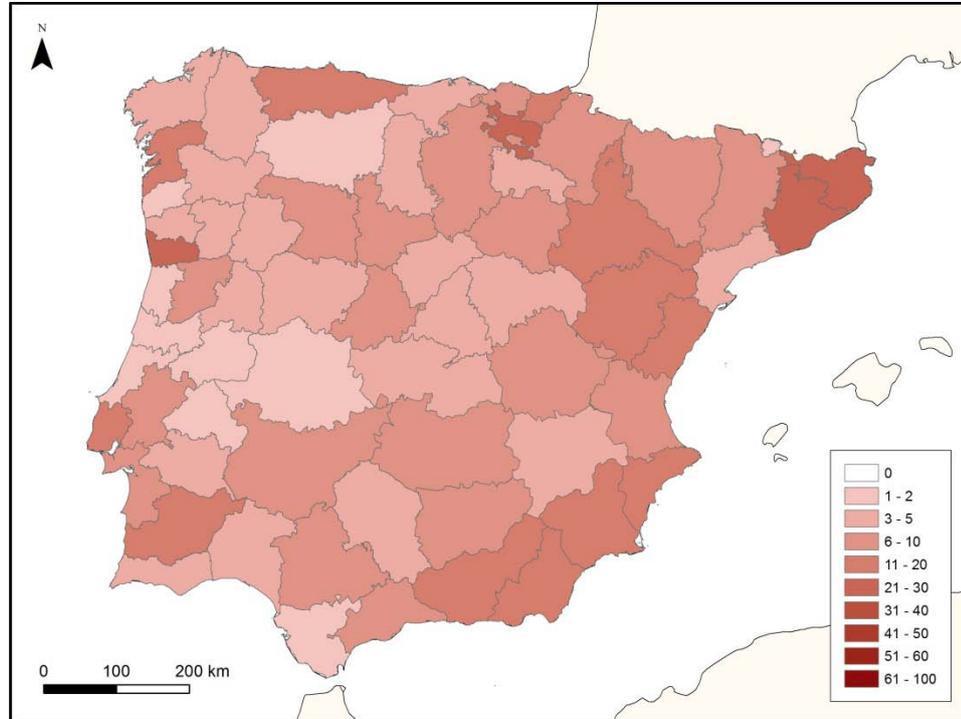


Figura 9 - Número de yacimientos con dataciones C14 por provincias españolas y distritos portugueses en 1978.

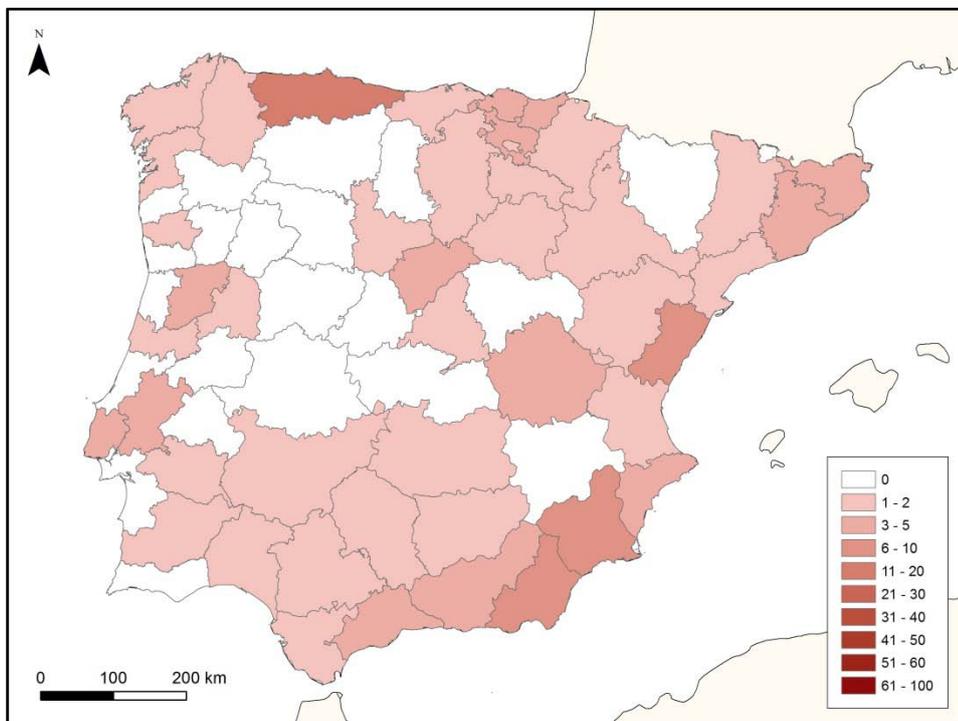


Figura 10 - Número de yacimientos con dataciones C14 por provincias españolas y distritos portugueses en 1996.

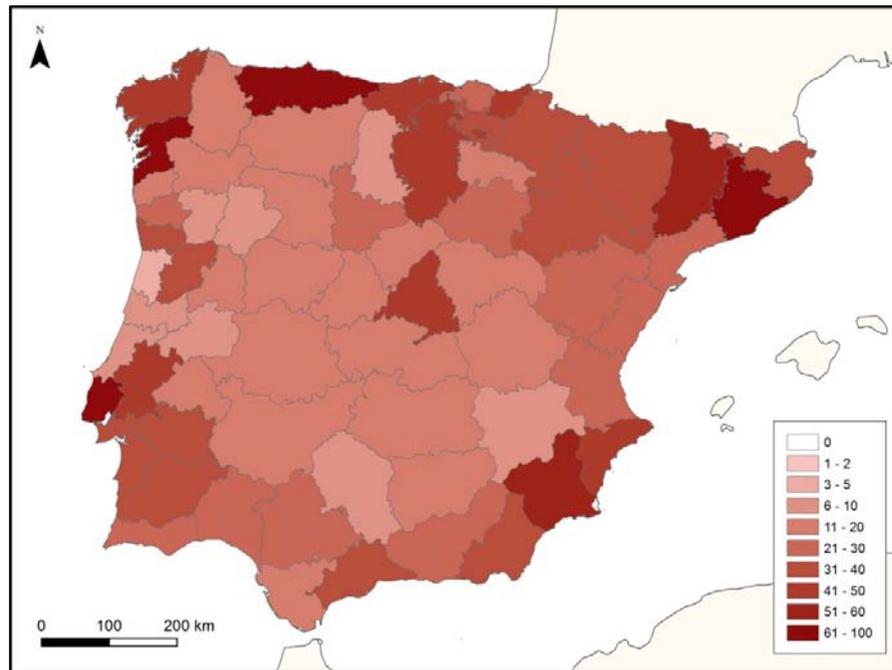


Figura 11 - Número de yacimientos con dataciones C14 por provincias españolas y distritos portugueses en 2016.

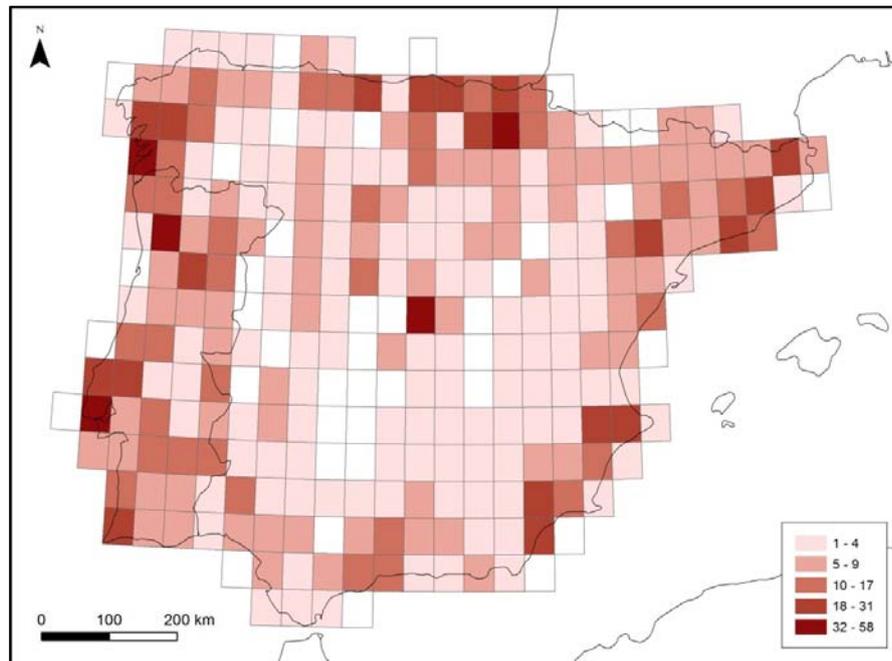


Figura 12 - Distribución del número de yacimientos con dataciones C14, agrupados según una cuadrícula geográfica de medio grado de paso.

La segunda perspectiva, la coherencia del conjunto de datos con hipótesis arqueológicas independientes, plantea algunos problemas adicionales, que quedan de manifiesto en la propia historia de la investigación y que tienen que ver, no tanto con la representatividad del conjunto de datos IDEArq-C14 con respecto al registro radiocarbónico peninsular, sino con la relación entre este último y la estructura cronológica real del registro arqueológico.

Esto queda de manifiesto si, en vez de ver la distribución de fechas en términos de demarcaciones administrativas, la representamos sobre una cuadrícula geográfica (en este caso se utiliza una malla de medio grado de paso) (Fig. 12). Si bien



todavía se aprecia el impacto del urbanismo sobre la incidencia del trabajo arqueológico y la importancia de ciertos intereses historiográficos, también son visibles ciertos vacíos que sólo pueden ser explicados por la falta de trabajo arqueológico: la mitad oriental de la provincia de Badajoz y occidental de Ciudad Real no son precisamente inhabitables (Fig. 4).

Ahora bien, como han demostrado Balsera y otros [Bal15], si utilizamos el conjunto de datos como base para un experimento estadístico basado en el análisis de la suma de probabilidades de las dataciones, encontramos que el registro recogido en IDEArq es compatible con la hipótesis más parsimoniosa sobre el proceso demográfico peninsular entre 7000 y 2000 cal BC: una curva logística, resultante de un crecimiento demográfico natural sostenido, con anomalías locales que pueden ser referidas a eventos históricos concretos. Esto significa que, considerado globalmente a escala general (tanto espacial como temporal), el conjunto de datos es una aproximación no sesgada a la estructura del registro arqueológico, y, por lo tanto, al propio registro radiocarbónico.

Si se acepta esta conclusión, podemos suponer que, en términos generales, muestras parciales obtenidas a partir del conjunto de datos podrían constituir una base observacional consistente para asesorar problemas arqueológicos que puedan ser formalizados en términos espacio-temporales a determinadas escalas.

Veamos lo que puede ser un ejemplo representativo de este tipo de estrategias. Durante el cuarto y tercer milenios cal BC, el rito de enterramiento dominante en la Península es colectivo. Los enterramientos colectivos tienen tres formas arqueológicas: tumbas megalíticas, cuevas sepulcrales naturales y cuevas sepulcrales artificiales (es decir, hipogeos). Aquí (Fig. 13) presentamos las probabilidades sumadas de fechas calibradas sobre huesos humanos de estos tres tipos de yacimiento. Estas distribuciones representan el uso de estos tipos de yacimientos como lugares de enterramiento. Lo más

evidente en estas tres distribuciones es la diferencia que existe entre cuevas sepulcrales, por una parte, e hipogeos y megalitos, por otra. Mientras que estos dos últimos coinciden en su distribución, la de las cuevas sepulcrales es más prolongada y multimodal. Que haya muertos del Neolítico antiguo y medio en cuevas sepulcrales y de época medieval en cuevas y megalitos no es sorprendente -son sitios buenos para esconder un cadáver-, pero el auge de muertos en cuevas durante el segundo milenio quizás requiera explicación. IDEArq facilita darle una dimensión espacial a este asunto. Si comparamos la distribución de los tres tipos de enterramiento en el cuarto milenio (Fig. 14) y el tercer milenio (Fig. 15) vemos que coinciden. Pero la distribución en el segundo milenio (Fig. 16) se concentra en el norte, con solo unos pocos casos sueltos en el resto de la Península. El rito colectivo sigue fuerte en el norte durante la Edad del Bronce cuando ha desaparecido prácticamente en el resto de la Península. Esta conclusión evidentemente no es original, pero el ejemplo ilustra la facilidad con la cual IDEArq permite jugar con el registro.

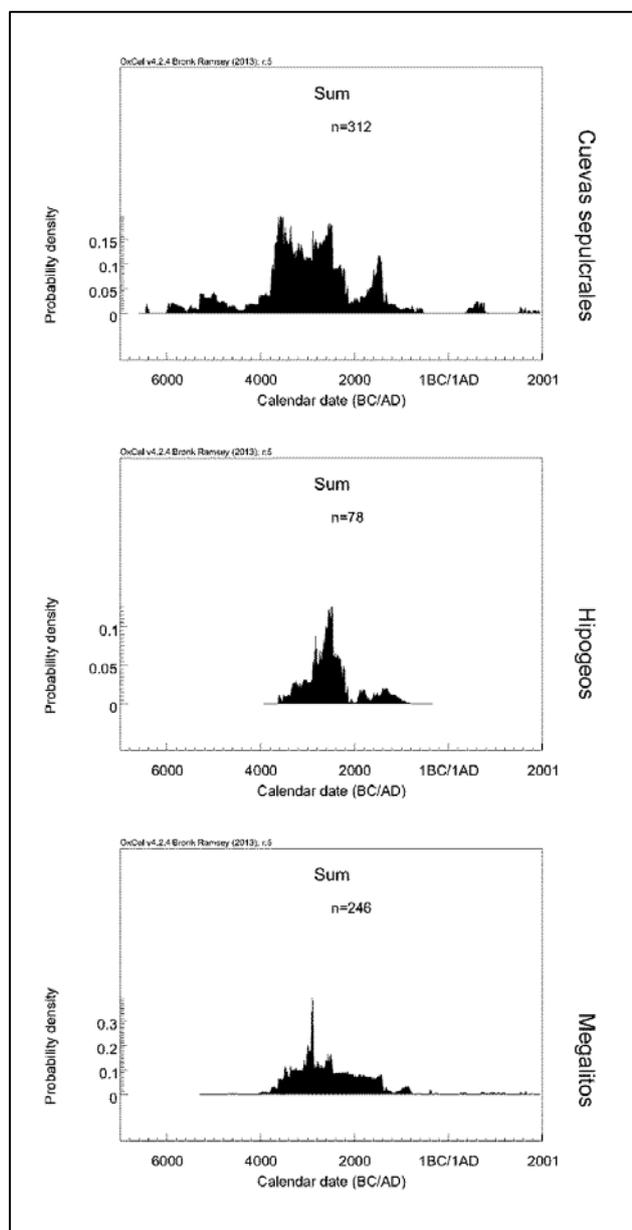


Figura 13 - Distribución de la suma de probabilidades de fechas calibradas sobre huesos humanos en tres tipos de yacimientos funerarios: cuevas sepulcrales, hipogeos y megalitos.

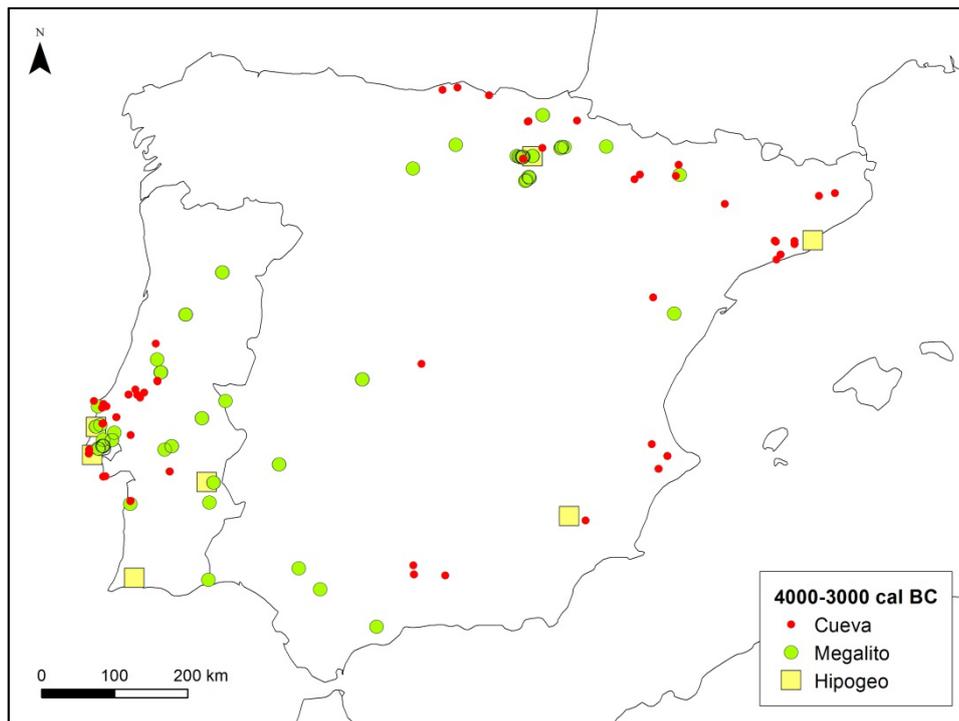


Figura 14 - Distribución de yacimientos de tipo funerario con dataciones C14 del cuarto milenio cal BC.

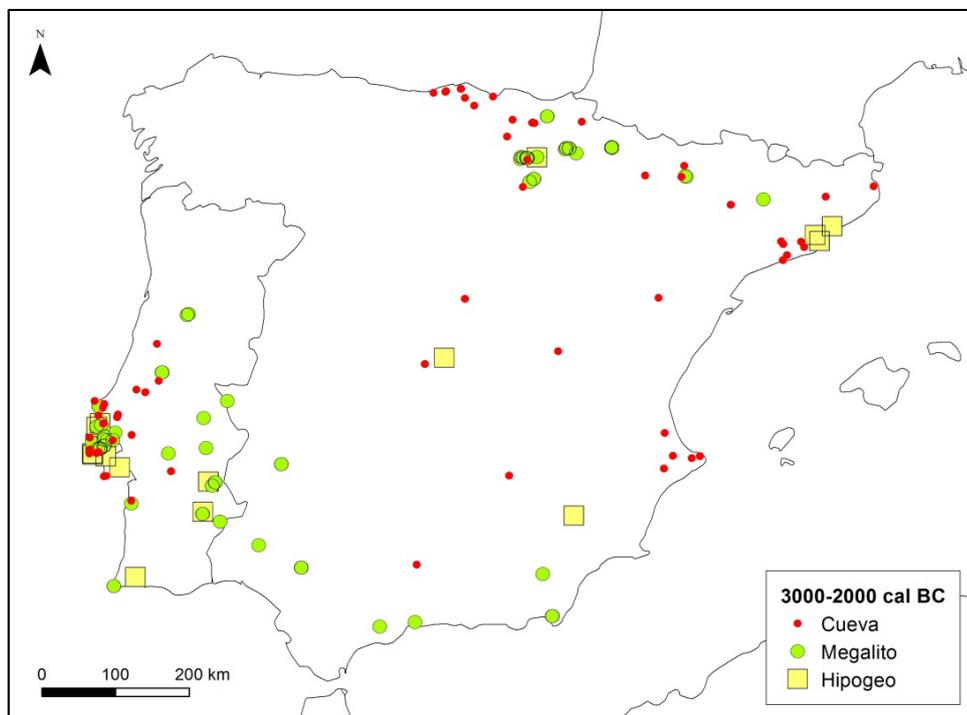


Figura 15 - Distribución de yacimientos de tipo funerario con dataciones C14 del tercer milenio cal BC.

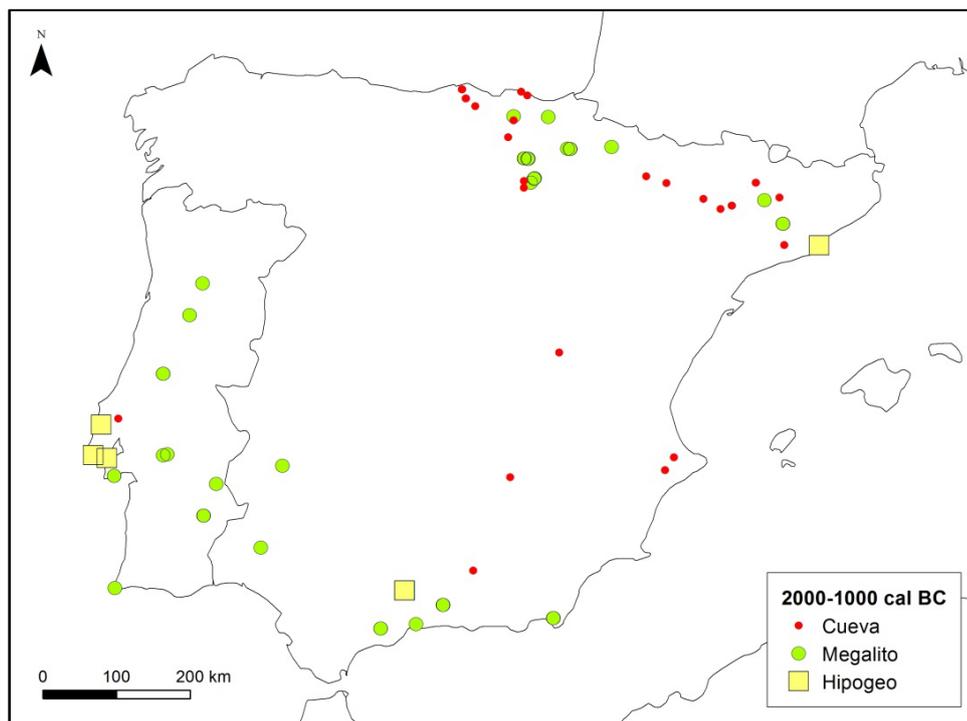


Figura 16 - Distribución de yacimientos de tipo funerario con dataciones C14 del segundo milenio cal BC.

El equipo editor confía en que su publicación en condiciones de libre acceso y su uso por la comunidad científica deberían de tener un efecto multiplicador en la disponibilidad de los datos cronométricos sobre la prehistoria ibérica. Esto se refiere, en primer término, a la introducción de correcciones y precisiones en los datos actualmente incorporados, a través de los mecanismos de comunicación entre usuarios y editores que incorpora el geoportal IDEArq. Pero también incluye la convicción de que la generalización de los criterios de interoperabilidad y acceso abierto que han inspirado el desarrollo del proyecto facilite la creación de redes de información científica abierta que hagan posible el acceso a un registro cronométrico normalizado y actualizado en tiempo real para la Prehistoria ibérica.

Bibliografía

- [Alm78] Almagro Gorbea M. y Fernández Miranda M. (eds.) (1978) C14 y prehistoria de la Península Ibérica. Madrid: Fundación Juan March.
- [Bal15] Balsera V., Díaz del Río P., Gilman A., Uriarte A., Vicent J.M. (2015) Approaching the demography of late prehistoric Iberia through summed calibrated date probability distributions (7000-2000 cal BC), *Quaternary Research*, Vol 386, pp. 208-211.
- [Bos16] Del Bosque González I. y Vicent García J.M. (2016) Presentación de IDEArq, Infraestructura de Datos Espaciales de investigación arqueológica, *Trabajos de Prehistoria*, Vol. 73, Nr 1, 2016, pp. 180-192.
- [Cas96] Castro Martínez P.V., Lull V., 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. Oxford: Archaeopress.
- [Cru09] Cruz Berrocal M., Gil-Carles Esteban J.M., Gil Esteban M., Martínez Navarrete M.I. (2005) Martín Almagro Basch, Fernando Gil Carles y el Corpus de Arte Rupestre Levantino, *Trabajos de Prehistoria*, Vol. 62, Nr 1, 2005, pp. 27-45.
- [Fer12] Fernández Freire C., Pérez Asensio E., Del Bosque González I., Uriarte González A., Vicent García J.M., Fraguas Bravo A., Parceró Oubiña C., Fábrega Álvarez P. (2012) Proposal for a Cultural Heritage Application Schema within the INSPIRE Directive, en: J.Gensel, D. Josselin, D.Vandenbroucke (eds.) *Multidisciplinary research on geographical information in Europe and beyond. Proceedings of the AGILE'2012 International Conference on Geographic Information Science, Avignon, April, 24-27, 2012*. Pp. 339-342.
- [Fer13] Fernández Freire C., Del Bosque González I., Vicent García J.M., Pérez Asensio E., Fraguas Bravo A., Uriarte González A., Fábrega Álvarez P., Parceró Oubiña C. (2013) A Cultural Heritage Application Schema: Achieving Interoperability of Cultural Heritage Data in INSPIRE, *International Journal of Spatial Data Infrastructures Research*, Vol 8, pp. 74-97.



- [Fer14] Fernández Freire C., Parcero Oubiña C., Uriarte González A. (eds.) (2014) A data model for Cultural Heritage within INSPIRE. *Cadernos de Arqueoloxía e Patrimonio* 35. Santiago de Compostela: CSIC.
- [Fer15] Fernández Freire C., Salas Tovar E., Gutiérrez González R., Vicent García J.M. (2015) A SDI's in Archaeology in the Iberian Peninsula: Achieving interoperability of Cultural Heritage data in INSPIRE, en: *AGILE'2015. International Conference on Geographic Information Science, 9-12 June 2015 - Lisbon, Portugal*.
- [Gil92] Gilman A. (1992) The Iberian Peninsula, 6000-1500 B.C., en: Enrich R.W. (ed.) *Chronologies in Old World Archaeology*. Chicago: University of Chicago Press. Vol. 1, pp. 295-301; Vol. 2, pp. 238-256.
- [Ins07] 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), *Official Journal of the European Union*, L108/1.
- [Ins10] INSPIRE Thematic Working Group Protected sites (2010) D.2.8.1.9 INSPIRE Data Specification on Protected Sites Guidelines v3.1. European Commission Joint Research Center.
- [Kuh62] Kuhn T. (1962) *The Structure of Scientific Revolutions*. University of Chicago Press: Chicago.
- [Lil16] Lilliois K.T., Blanco-González A., Drake B.L., López-Sáez J.A. (2016) Mid-late Holocene climate, demography, and cultural dynamics in Iberia: A multi-proxy approach, *Quaternary Science Reviews*, Vol. 135, pp. 138-153.
- [Par13] Parcero Oubiña C., Fábrega Álvarez P., Vicent García J.M., Uriarte González A., Fraguas Bravo A., Del Bosque González I., Fernández Freire C., Pérez Asensio E. (2013) Conceptual basis for a cultural heritage data model for INSPIRE, *Revue Internationale de Géomatique / International Journal of Geomatics and Spatial Analysis*, Vol. 23, Nr 3-4, pp. 445-467.
- [Uri13] Uriarte González A., Parcero Oubiña C., Fraguas Bravo A., Fábrega Álvarez P., Vicent García J.M., Pérez Asensio E., Fernández Freire C., Del Bosque González I. (2013) Cultural Heritage Application Schema: a SDI framework within the Protected Sites INSPIRE Spatial Data Theme, en: G.Earl, T.Sly, A.Chrysanthi, P.Murrieta-Flores, C.Papadopoulos, I.Romanowska, D.Wheatley (eds.) *Archaeology in the Digital Era. Papers from the 40th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA), Southampton, 26-29 March 2012*. Amsterdam: Amsterdam University Press. Pp. 279-290.