

Buscar y Encontrar información radiométrica
en Internet.

Una herramienta de búsqueda global

Miquel Colobrán, Igor Bogdanovic, Juan A. Barceló

Buscar y encontrar información radiométrica en Internet. Una herramienta de búsqueda global

Miquel Colobrán, Igor Bogdanovic, Juan A. Barceló

Universitat Autònoma de Barcelona

RESUMEN

En este trabajo se exponen los fundamentos del proyecto GoGet, un motor de búsqueda en Internet capaz de explorar en diferentes bases de datos radiométricas online y producir listados exhaustivos de dataciones isotópicas y sus contextos arqueológicos. La herramienta permite la búsqueda tanto por medio de interrogaciones temporales (dataciones dentro de determinado intervalo temporal) como geográficas (dataciones de muestras procedentes de determinada área geográfica). GoGet proporciona el listado de las diferentes bases de datos online que contienen esa información precisa, y permite editar las repeticiones (cuando una misma datación aparece en más de una base de datos).

Palabras clave: radiocarbon, database, data formalization, format standadization, search engine, Internet

1 Introducción

La dispersión y escasa congruencia de la información es una característica muy típica de los datos disponibles en internet. La cantidad de información accesible en la world wide web es inmensa, sin embargo, es tan dispersa que resulta muy difícil acceder a ella en la práctica. Si por información entendemos un conjunto organizado de datos procesados, que constituyen un mensaje que cambia el estado de conocimiento del sujeto o sistema que recibe dicho mensaje, internet no contiene “información”, sino datos pendientes de procesar y pendientes de ser incluidos en un mensaje.

La información radiométrica es imprescindible para la investigación histórica. En la actualidad está dispersa en multitud de repositorios, está expresada en una gran diversidad de formatos y de idiomas, de manera tal que resulta muy difícil el intercambio de datos. La forma más corriente en Internet es a través de bases de datos online que contienen información acerca de dataciones radiométricas. Esas bases de datos, sin embargo, son muy distintas entre sí. Algunas son internacionales en el sentido que contienen muestras datadas de lugares muy diferentes; contienen numerosos datos, pero no son exhaustivas y la información referida suele ser muy básica. Otras contienen datos de regiones específicas o bien son muy concretas en términos del intervalo de tiempo o el periodo histórico cubierto. Incluso algunas son resultado del trabajo de un investigador o grupo.

Lo que se necesita es una herramienta que no sólo sea capaz de buscar en repositorios distintos, sino que traduzca esa diversidad de formatos, lenguajes y modelos representacionales en una única ficha característica que estandarice el proceso de consulta. El propósito de GoGet es por tanto la integración de las bases de datos isotópicas, pero respetando las especificidades de la información contenida en cada una de las bases de datos exploradas. Proporciona un punto de acceso común, disponible online, y fácilmente accesible para toda la comunidad investigadora y académica.

Goget es nuestra propuesta. Constituye una aplicación práctica de la propuesta de “telearqueología” (Bogdanović 2002; Barceló, Bogdanović, Vicente 2004; Barceló,

Bogdanović, Piqué 2004), que se concibe como un conjunto de conceptos computacionales, soluciones y herramientas para la conexión, integración y manipulación de datos heterogéneos distribuidos en la World Wide Web, y que tiene como objetivo la generación de conocimiento arqueológico de forma autosuficiente. En esta nube ampliamente abierta de conocimiento arqueológico, la idea de "publicación" y construcción del conocimiento adquiere un nuevo sentido. La telearqueología es un marco abierto para discutir la publicación de datos, teorías e interpretaciones de la manera más eficiente a través de la formalización y sistematización de conjuntos de datos, y mediante el uso de interfaces de interoperabilidad y conexiones a diferentes bases de datos u otros contenidos en la WEB.

2 Los datos radiométricos

Aunque no existe un estándar descriptivo universal, la mayoría de bases de datos online definen los atributos que describen cada uno de los registros en términos de la información mínima que se debiera incluir poder evaluar la fiabilidad de la información temporal medida con un reloj isotópico (Kra, 1988) (Levine & Stanish, 2014) (Millard, 2014) (Jull & McNichol, 2014) (Wood, 2015). De acuerdo a esas indicaciones, los atributos mínimos que definirán el suceso isotópico debieran ser:

- **CÓDIGO:** El número de código individual del laboratorio, que se antepone a las mediciones de radiocarbono de ese laboratorio en particular.
- **NUMERO DE REFERENCIA** de la muestra analizada, según el sistema de catalogación de los materiales procedentes del yacimiento arqueológico estudiado
- **ERROR DE LABORATORIO** (Scott et al., 2007).
- **EDAD CONVENCIONAL ¹⁴C.** Resultado de la medición isotópica, en años antes del presente (bp). El uso de la terminología “Edad de Radiocarbono Convencional” en un informe indica que el resultado ha sido corregido en función del fraccionamiento isotópico.
- **DESVIACIÓN TÍPICA DE LA EDAD CONVENCIONAL ¹⁴C.**
- **METODO DE MEDIDA ¹⁴C:** AMS, CENTELLEO, etc. Cf. Capítulo 4 acerca de los distintos tipos de medida.
- **PORCENTAJE CARBONO MARINO.**
- **EFFECTO RESERVORIO:** Método usado para determinar el valor de corrección de reservorio marino para muestras afectadas por carbono de procedencia marina. Hay dos posibles correcciones aplicables a una fecha de radiocarbono para carbonatos marinos: la corrección localizada de yacimientos (Delta±R) y la corrección global. La corrección global aplica automáticamente a través del programa de calibración, basándose en la edad de la muestra. Esta corrección oscila entre -200 y 500 años. Estos valores tienen en cuenta que se necesitan entre 200 y 500 años para que el dióxido de carbono presente en la atmósfera se incorpore y distribuya (equilibre) a través de la columna de agua del océano.
- **Delta±R.** Valor local del efecto reservorio. Este valor varía de un lugar a otro en función del afloramiento localizado, la escorrentía de agua dulce y otras condiciones relacionadas con la profundidad, la circulación, la salinidad, etc.
- **EDAD ¹⁴C CORREGIDA POR IMPACTO EFFECTO RESERVORIO.** Se aplica una corrección Delta±R a una fecha de radiocarbono que ya ha sido corregida con la corrección global del yacimiento marino. El valor aportado por el cliente se resta o se suma a esta antigüedad ya corregida (dependiendo de si se trata de un valor Delta+R o Delta-R)
- **EDAD ¹⁴C CALIBRADA.**
- **CURVA DE CALIBRACIÓN EMPLEADA** (Cf. Capítulo 4 bis)

No existe en la actualidad un formato específico para codificar la edad ¹⁴C calibrada. No se trata de un escalar (valor numérico), sino de un intervalo de probabilidad. Programas de calibración como OxCal han generalizado un formato tabular para representarlo, si bien esa tabla se divide

de manera diferente para diferentes muestras, según sea la probabilidad de los sub-intervalos dentro del intervalo de calibración definido por 2 veces la desviación típica. En este orden de cosas, la manera más simple de representar el intervalo de confianza de una edad ^{14}C calibrada implica el uso de dos atributos, cada uno definido por un único valor numérico: extremo inferior/extremo superior del intervalo de confianza de la edad calibrada, especificando si se trata del intervalo definido por una desviación típica (68% de confianza) o 2 desviaciones típicas (95% de confianza). Es importante utilizar un formato de fecha normalizado, es decir, año calendárico astronómico (con año 0), y usando valores negativos para “años antes de nuestra era” y valores positivos para “años de nuestra era”. Cualquier otra forma de codificar las posiciones temporales en la escala calendárica daría problemas a la hora de combinar la información.

Estos atributos constituyen la información mínima necesaria para representar la información cronométrica de un registro individual (muestra datada). Pero en la mayoría de los casos, no constituyen una información suficiente. Debemos caracterizar la muestra que se ha analizado, porque sólo con esa información podremos validar la datación misma. Por ese motivo, a los atributos anteriores debíamos añadir:

- TIPO DE MUESTRA ANALIZADA: Carbón, madera semilla, hueso animal, hueso humano, etc.
- TAXONOMÍA MUESTRA ANALIZADA: Por ejemplo, *triticum diccocom*, *robinia pseudoacacia*, *cervus elaphus.*, *homo sapiens*, etc.
- ALTERACIONES DE LA MUESTRA: quemada/no quemada en origen
- EDAD RELATIVA DE LA MUESTRA. Vida media del individuo del cual procede la muestra en el momento de su muerte.

Además, debíamos consignar también:

- $\delta^{13}\text{C}$. Valor del fraccionamiento isotópico de la muestra.
- METODO DE MEDIDA $\delta^{13}\text{C}$.: Método usado para determinar fraccionamiento isotópico de la muestra: Medido, estimado.
- ERROR STANDARD $\delta^{13}\text{C}$.
- $\delta^{15}\text{N}$. Determinación de la proporción $^{15}\text{N}/^{14}\text{N}$ del total de nitrógeno en la muestra.

Cada vez se utilizan más las relaciones isotópicas estables de nitrógeno ($\delta^{15}\text{N}$) y carbono ($\delta^{13}\text{C}$) para cuantificar la longitud de la cadena alimentaria y la anchura de los recursos tróficos respectivamente. A lo largo de una serie de longitudes conocidas de la cadena alimentaria, examinamos cómo cambiaron $\delta^{15}\text{N}$ y $\delta^{13}\text{C}$ dentro de las cadenas alimentarias y entre ellas. Estos valores proporcionan información importante sobre el entorno del que procede la muestra o las mezclas de materiales utilizados para producirla, ya que el valor isotópico de la muestra refleja la composición isotópica del entorno inmediato.

2.1. Inferencia cronológica.

Dado que la inferencia cronológica es, ante todo, una operación combinatoria de dataciones individuales, debemos introducir aquella información que nos permita definir conjuntos homogéneos. Esos atributos son espaciales y “crono-culturales”. En el primer caso, introduciremos la referencia al yacimiento arqueológico del cual procede la muestra analizada, así como las coordenadas UTM del yacimiento y las coordenadas espaciales locales (cartesianas) de la localización precisa de la muestra dentro del yacimiento¹. En el segundo

¹ El uso de coordenadas geográficas universales (UTM) es muy engorroso para determinar la localización precisa de los sucesos depositacionales o unidades espaciales mínimas: una medida en coordenadas geográficas universales (UTM) con precisión de 1 metro requiere seis dígitos; ocho dígitos si la precisión debe ser centimétrica. Cuantos más decimales requiera la coordenada espacial del suceso depositacional, más probables son los errores de proceso. En el caso de Galileo, la precisión de posicionamiento horizontal de la señal abierta al 95%, para un receptor de doble frecuencia, es de 4 m (8 m para la precisión vertical). (Fuente: ESA. Navipedia https://gsse.esa.int/navipedia/index.php/Galileo_Performances. Consultado el 8 de Junio de 2019). Estos instrumentos triangulan su posición utilizando su distancia de al menos cuatro satélites. Debido a que miden esta distancia basándose en el tiempo que tarda una señal de satélite en llegar a ellos, incluso los errores más pequeños -hasta unas pocas mil millonésimas de segundo- pueden afectar negativamente a la precisión. Los errores en la posición de la órbita del satélite pueden llevar a una pérdida de precisión de alrededor de 2,5 metros. Los errores de reloj de los satélites pueden añadir otros 1,5 metros. Y las perturbaciones en la troposfera y la ionosfera pueden sumar otros uno y cinco metros respectivamente, aún más si el satélite está cerca del horizonte o durante periodos de intensa actividad solar. El mayor error se debe, con mucho, a los efectos multitrayecto, en los que las señales de

caso, se trata de introducir una etiqueta conceptual que permita combinar dataciones del mismo “periodo histórico”. Si bien en la práctica esa referencia suele ser subjetiva y se deja la asignación crono-cultural del investigador o investigadora que ha encargado y publicado la datación, sería recomendable formalizar la definición de dicho “periodo” siguiendo los protocolos *ChronOntology* (<https://chronontology.dainst.org/>) o *PeriodO* (<http://perio.do/en/>).

Obviamente, una base de datos cronométrica se crea con el fin de estimar la posición temporal de sucesos arqueológicos. Por consiguiente, la consulta de la base de datos estará marcada por la definición del suceso arqueológico que se pretende datar. Los sucesos se ordenan del más específico – suceso isotópico – al más genérico, -periodo histórico-, definido este último como conjunto uniformemente distribuido de sucesos históricos. En teoría, el usuario de la base de datos introducirá como consulta una definición formalizada del suceso cuya posición temporal en la escala calendárica se pretende conocer y el sistema devolverá el conjunto de dataciones asociadas a sucesos que el sistema conoce que son contemporáneos al interrogado. La calidad de esa respuesta depende de la manera como está organizada la información. Las bases de datos presentadas en los capítulos anteriores no son relacionales, sino que los distintos registros son independientes y se relacionan en una única tabla (base de datos plana o fat file database).

Dadas las características de estas bases de datos, las consultas se reducen a listados por atributos o por combinación de atributos. Por ejemplo:

- Todas las dataciones del yacimiento *Y*
- todas las dataciones de todos los yacimientos cuyo extremo inferior de la edad calibrada se sitúa entre -2000 y -4000.
- todas las dataciones cuya diferencia entre el extremo inferior y el extremo superior de la edad calibrada se sitúe entre -2000 y -4000.
- todas las dataciones de sólo aquellos yacimientos del área geográfica definida por las coordenadas *XXXYYYY*, cuya diferencia entre el extremo inferior y el extremo superior de la edad calibrada se sitúe entre -2000 y -4000
- Sólo aquellas dataciones sobre hueso humano de sólo aquellos yacimientos del área geográfica definida por las coordenadas *XXXYYYY*, cuya diferencia entre el extremo inferior y el extremo superior de la edad calibrada se sitúe entre -2000 y -4000
- Sólo aquellas dataciones sobre semilla de sólo aquellos yacimientos del área geográfica definida por las coordenadas *XXXYYYY*, cuya diferencia entre el extremo inferior y el extremo superior de la edad calibrada se sitúe entre -2000 y -4000 y que estén asociadas al periodo “Campaniforme”

3 Información radiométrica en Internet

3.1. Bases de datos cronométricas en Internet.

satélite llegan al receptor por trayectorias múltiples o indirectas, por ejemplo, al rebotar en paredes de edificios en cañones urbanos. En condiciones de cielo abierto, una coordenada espacial así calculada tiene una precisión de aproximadamente dos metros. Los receptores satelitales avanzados mejoran drásticamente la precisión espacial utilizando datos de corrección para eliminar los errores. Una forma de obtener estos datos consiste en monitorizar las señales procedentes de la red satelital desde una estación base en una ubicación conocida. Las desviaciones de la posición de la estación base se observan y se envían a un rover - un vehículo tripulado o no tripulado equipado con un receptor satelital - lo que le permite obtener una lectura más precisa de la posición. En condiciones favorables, este enfoque puede ser utilizado para lograr una precisión de nivel de centímetros, siempre y cuando la estación base y el rover no estén demasiado alejados (Morris y Conner, 2017). La localización de la muestra analizada dentro la unidad espacial mínima de referencia debiera especificarse, por tanto, en un sistema cartesiano local (*x,y,z*), debido a la escasa resolución de la mayoría de receptores satelitales (GPS, Galileo). De este modo, cada muestra analizada, que representa un suceso isotópico debe poder asociarse de manera no equívoca a un poliedro cerrado que conforme una partición del espacio [excavado]; cada punto en el espacio arqueológico, datado o no datado, debiera pertenecer a una unidad espacial y sólo a una, sólo así podremos reconstruir la complejidad de las relaciones espacio-temporales en el sitio arqueológico (cf. Capítulo 9).

La explosión de la web ha permitido el acceso a todo tipo de datos a cualquier persona que tenga acceso a un ordenador o dispositivo móvil (tablets, smartphone, tv inteligente, etc) conectado a Internet. Las bases de datos, por ello, se han visto muy beneficiadas dado que muchas se han convertido en bases de datos online, de consulta a través de un navegador. Como ejemplo tres de las más conocidas, Radon (<https://radon.ufg.uni-kiel.de/>), Archaeological Data Service (https://archaeologydataservice.ac.uk/archives/view/c14_cba/) and Canadian Archaeological Radiocarbon Database (<https://www.canadianarchaeology.ca/>). Cada una conteniendo varios miles de dataciones.

3.2. Carencias de las bases de datos cronométricas online

Muchas bases de datos permiten buscar y localizar la información cronométrica que contienen, esencialmente, una descripción básica de ciertos sucesos isotópicos. Sin embargo, no son capaces de realizar inferencias cronológicas porque no hay manera de interrogar al sistema acerca de la posición temporal en la escala calendárica de un suceso arqueológico.

Para que un sistema de información cronométrica pueda funcionar como un sistema experto capaz de estimar la “datación” más probable de un suceso arqueológico cuya temporalidad es desconocida, es preciso definir un algoritmo capaz de encontrar en la base de datos un registro “contemporáneo” a la descripción del suceso que se investiga. Si y sólo si existe en la base de datos un registro temporal (uno o varios sucesos isotópicos) suficientemente contemporáneos, la posición temporal del suceso conocido será transferida al suceso que se investiga y se ejecutará un cálculo de la validez de esa estimación temporal.

Tal y como se ha argumentado en el capítulo 11, dos son las maneras de calcular la “contemporaneidad” entre dos sucesos en el caso de no disponer de una medición isotópica para ambos: 1) que estén en una misma localización espacial, 2) que ambos hayan sido asignados al mismo “periodo histórico”.

En el primer caso, nos estamos refiriendo al suceso depositacional, que se puede definir en términos de la descomposición geométrica en 3D del espacio arqueológico (cf. Capítulo 9): toda evidencia encontrada dentro de una misma unidad espacial de referencia tendrá la misma posición temporal en la escala calendárica, dentro de un intervalo de error definido en términos de la duración de dicho suceso depositacional. Por consiguiente, en ausencia de cualquier otra información temporal, el sistema de información debiera ser capaz de usar información espacial en tres dimensiones y relaciones estratigráficas para poder combinar registros. Los registros en el sistema de información debieran conectar, no sólo una descripción suficiente de la unidad espacial mínima a la que pertenecen, sino a las relaciones estratigráficas que mantienen entre sí (p.e. matrices de Harris ampliadas).

Distinta evidencia arqueológica será contemporánea, en ausencia de información espacial y estratigráfica relevante, si ambas pueden ser asignadas a un mismo intervalo temporal cuyo momento de inicio y final han sido definidos objetivamente, y dentro del cual cierta categoría material determinada –o rasgo cultural- es aproximadamente constante. Según esto, un periodo histórico coincide con el periodo de vigencia de la característica en cuestión: el intervalo temporal en el que la ocurrencia de dicha característica tiene una probabilidad no nula (cf. Capítulos 9 y 11).

Para que un sistema de información sea capaz de realizar este tipo de inferencias, debemos añadir a los registros ya conocidos –sucesos isotópicos datados- nuevos registros que, aún cuando no estén datados por métodos isotópicos, mantengan relación de contemporaneidad con los sucesos isotópicos datados. Ello conlleva un replanteamiento sobre la representación de estos datos y sus relaciones, que requiere un modelo que no puede ser representado en un flat file y que solamente puede ser implementado en una base de datos completa.

3.3. Motores de búsqueda en Internet.

La gran cantidad de información existente en Internet ha traído consigo un problema. ¿Dónde encontrar la información que necesito? Con millones de web potencialmente consultables, el filtrado de ellas a fin de conocer las relevantes a la búsqueda se reveló como un problema de gran

envergadura. Así pues, los motores de búsqueda fueron el siguiente paso lógico. Un motor de búsqueda (*search engine*) es un sistema capaz de recoger información de diferentes webs, indexarla y organizarla a fin de que el usuario pueda encontrarla sin navegar por todas las web existentes. El ejemplo paradigmático lo constituye Google.

Los motores de búsqueda cuentan con rastreadores muy potentes que navegan y analizan los sitios web y crean una base de datos con la información recolectada. No contienen la información en sí, sino solo un pequeño subconjunto y la dirección URL del origen de la información obtenida. De ese modo el usuario, tras la consulta (generalmente con palabras clave) puede obtener una relación de páginas web que contienen los términos y visitar las web que contiene toda información. Los elementos básicos (figura 1) de un motor de búsqueda (Levene, 2010) son:

- Crawler (rastreador). Este elemento es un programa que “visita sitios web”, lee las páginas y extrae la información mínima necesaria para añadir a la base de datos del motor. Detectan los enlaces, por lo que tienen la capacidad de seguirlos a otras páginas del sitio web hasta leerlas todas. Los crawlers también se conocen como spider (araña), web bot, Internet bot o bot.
- Indexador. Es el elemento que pone la información recogida del crawler en la base de datos del motor de búsqueda, de modo que la página pueda aparecer en los resultados de búsqueda.
- Base de datos: Es el corazón de la arquitectura. Permite el almacenado de la información de la página visitada así como la rápida recuperación de estos ante una consulta.
- Interfaz de usuario: Es el elemento que se comunica con el usuario y con la base de datos del motor de búsqueda. Casi siempre es un servidor web. Se encarga de crear una consulta a partir de las palabras clave o información creada por el usuario. Seguidamente realiza la consulta sobre la base de datos y muestra en el navegador del usuario el resultado.

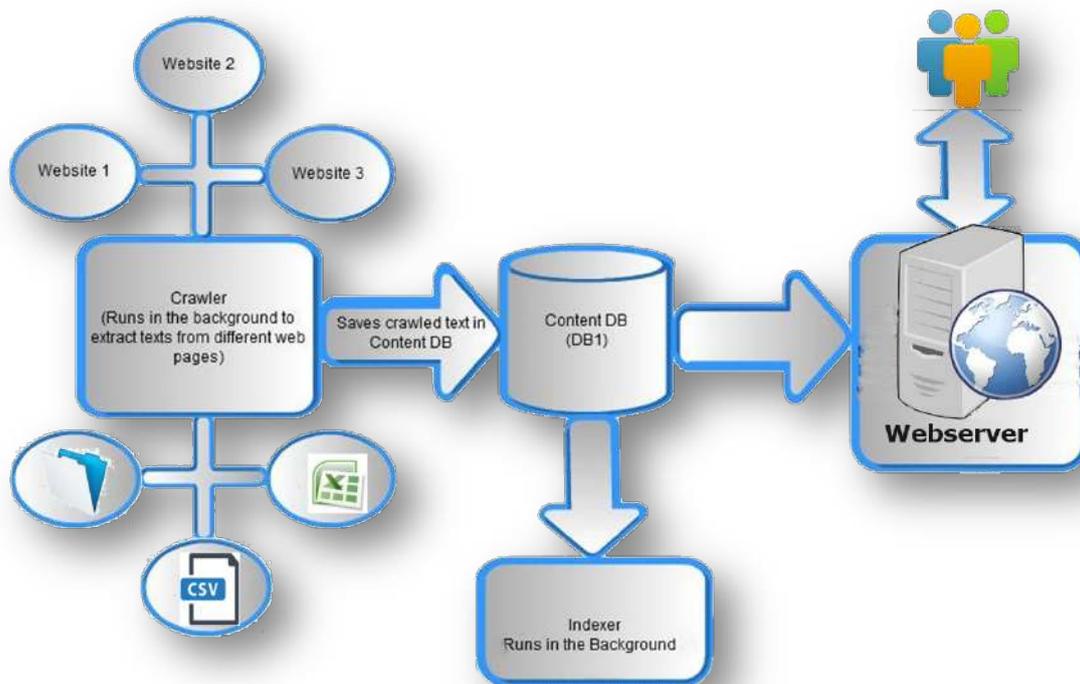


Figura 1. Arquitectura básica de un motor de búsqueda

4. Motor de búsqueda de dataciones radiocarbónicas: Goget

4.1. Estructura de Goget

GoGet es un rastreador web de dataciones de ^{14}C . Su funcionamiento es, por tanto, semejante al del buscador Google. Por un lado, cuenta con su propio buscador por texto, que compara las palabras introducidas con las existentes en su base de datos interna. Por otro lado, contiene un formulario en donde se definen las características que queremos que tengan las dataciones a mostrar, acotando así los resultados de la búsqueda. La página principal de GoGet es accesible desde la url

<http://www.ibercrono.org/goget/>

GoGet no pretende contener toda la información cronométrica de las muestras o contextos arqueológicos datados, sino indexar la información básica existente en distintas bases de datos regionales, nacionales o internacionales. Lo que hace el programa es indexar la información temporal contenida en las bases de datos. Al igual que un motor de búsqueda, como Google, GoGet pregunta al usuario por los términos y busca la información. Como resultado, GoGet muestra una lista de fechas que coinciden con los criterios. Toda la información de la datación se encuentra en la base de datos original, a la que GoGet redirecciona. Por consiguiente, GoGet no es, en realidad un repositorio o base de dataciones, sino un buscador de bases de datos que contienen dataciones, y que muestra como resultado de las distintas operaciones de búsqueda la información contenida en la base de datos original

GoGet sigue la estructura y contiene los elementos de cualquier motor de búsqueda. Por ello su funcionamiento es muy similar al de Google. Se descompone en tres capas (figura 2). La parte inferior (capa de datos) consiste en información completa de datos de radiocarbono, georeferenciación espacial global a escala geográfica, georeferenciación espacial local a microescala (dentro del sitio) e información de contexto. Consta de las bases de datos existentes en Internet, y por ello heterogéneas con información diferente y almacenada de manera

diferente. Goget solo lee esta información usando la segunda capa, llamada plataforma de interoperabilidad, y la deposita en la tercera capa, la base de datos. En ella, junto con una aplicación web es donde el usuario puede hacer sus consultas.



Figura 2. Capas (layers) del motor de búsqueda Goget

4.2. Funcionamiento de Goget

4.2.1. Capa de datos

Se considera en esta capa todas las bases de datos existentes en Internet, y por ello heterogéneas con información diferente y almacenada de manera diferente. Estas bases de datos están hechas y mantenidas por aquellos investigadores y grupos de investigación que las ponen a disposición de la comunidad radiocarbónica a través de Internet. Nuestro objetivo es que estas bases de datos no realicen ningún cambio en su modo de funcionamiento y por lo tanto, el ‘trabajo’ de obtener la información que contienen reca en la segunda capa.

4.2.2. Capa de interoperabilidad

Goget lee la información en la segunda capa, llamada plataforma de interoperabilidad. Esta está construida con un conjunto de agregadores (elementos de software o middleware) que son capaces de leer las bases de datos y "traducir" la información para que pueda ser almacenada posteriormente en la base de datos. Los agregadores, dado que las bases de datos de radiocarbono son extremadamente heterogéneas, necesitan de un agregador hecho a medida para cada una. De hecho lo que realmente hacen es ‘leer’ toda la información de la base de datos que se encuentra disponible en Internet. De ese modo el creador no debe realizar cambios en su base de datos. Estos agregadores están realizados en el lenguaje de programación Python. Periódicamente se vuelve a descargar la información de la base de datos y se incluyen las modificaciones en Goget.

Toda la información obtenida de la base de datos por el agregador se convierte a un formato interno común para todas las bases de datos ‘visitadas’. El formato interno es muy parecido a XML y los agregadores usan este formato para comunicarse con la capa superior. Este XML se realiza de forma automática. Así, cualquier base de datos externa podría crear un archivo que mantenga esta estructura. Básicamente la estructura del archivo xml tiene tres "bloques de datos": la parte obligatoria, la opcional y la personalizada. La base de este formato XML para la base de datos de radiocarbono es una ontología que define los conceptos, atributos y relaciones entre conceptos (figura 3).

XML es la abreviatura de eXtensible Markup Language. Fue diseñado para almacenar y transportar datos con la intención de ser legible tanto por humanos como por máquinas. Para hacer posible la comunicación con todos los sistemas informáticos, XML es un formato de texto simple y muy flexible derivado de SGML (ISO 8879). En la Figura se muestra un ejemplo de un archivo XML que contiene datos de radiocarbono.

```

<RADIOCARBON>
  <DATE>
    <LAB_CODE> OXAR2000 </LAB_CODE>
    <LABORATORY> OXA </LABORATORY>
    <DESCRIPTION> Carbo en estructura </DESCRIPTION>
    <AGE> 5800 </AGE>
    <YEAR> 1985 </YEAR>
    <COORDS> LAT LONG </COORDS>
    <LATITUDE> 0.25148 </LATITUDE>
    <LONGITUDE> 30.2154 </LONGITUDE>
    <HUSO> </HUSO>
    <ZONE> </ZONE>
    <DATUM> </DATUM>
    <AUTHOR> Laboratorio de arqueología cuantitativa </AUTHOR>
    <CONTACT> equ@unq.edu </CONTACT>

    <METHOD> AMS </METHOD>
    <DESVIATION> 30 </DESVIATION>
    <C13> 1 </C13>
    <C13_C12_RATIO> </C13_C12_RATIO>
    <COMMENTS> </COMMENTS>
  </DATE>
  <DATE>
    <LAB_CODE> Beta42515 </LAB_CODE>
    <LABORATORY> Beta </LABORATORY>
    <DESCRIPTION> Beta </DESCRIPTION>
    <AGE> 1000 </AGE>
    <YEAR> 198 </YEAR>
    <COORDS> 0 0 </COORDS>
    <LATITUDE> </LATITUDE>
    <LONGITUDE> </LONGITUDE>
    <HUSO> </HUSO>
    <ZONE> </ZONE>
    <DATUM> </DATUM>
    <AUTHOR> Laboratorio de arqueología cuantitativa </AUTHOR>
    <CONTACT> equ@unq.edu </CONTACT>

    <METHOD> </METHOD>
    <DESVIATION> </DESVIATION>
    <C13> </C13>
    <C13_C12_RATIO> </C13_C12_RATIO>
    <COMMENTS> </COMMENTS>
  </DATE>
</RADIOCARBON>

```

Figura 3. Información radiocarbónica de una datación expresada en XML

Algunas de las tareas que debe realizar un agregador, de forma adicional, es el control de errores, de duplicados, que la información mínima exista y esté correcta, que el formato del número de datación sea correcto, una primera gestión de duplicados, etc.

En esta segunda capa, cuando un agregador ha obtenido la información y la ha ‘traducido’ al formato interno, común para todas las bases de datos como se ha dicho, tiene que colocarla en la base de datos de Goget. Ello se lleva a cabo mediante otro elemento llamado cargador (*uploader*) que se encarga de poner la información dentro de la base de datos final. El cargador también está realizado en el lenguaje de programación Python y en este caso es único dado que todos los agregadores han realizado la traducción a un formato interno común.

4.2.3. Base de datos e interfaz de usuario web

La tercera capa, contiene dos subsistemas. Los cargadores, que transmiten la información (ya en formato estándar de Goget) a una ontología donde se realizan comprobaciones, extracción de características, detección de duplicidades, traducción de términos e incorporación de ellos en los diccionarios. Todo este proceso termina con la indexación. La información resultante queda almacenada dentro de la base de datos que servirá a la interfaz del usuario. Se realiza un informe resultado de todo este proceso con las anomalías detectadas como pueden ser duplicidades

dentro de la misma base de datos, dataciones incorrectas (las coordenadas no existen), etc. Este informe puede ser remitido al propietario de la base de datos a fin de mejorar la calidad de estos. Finalmente el portal web contiene la interfaz de usuario que, como cualquier motor de búsqueda, permite la consulta a través de un servidor web a la comunidad de investigadores de radiocarbono (figura 4). Así pues, el portal web contiene la estructura estándar de cualquier sitio web, una base de datos (realizada en Mysql) y un servidor web (realizado con tecnología php).

The screenshot shows the user interface of the GoGet website. At the top, there is a navigation bar with the GoGet logo, the title 'The Radiocarbon Searcher', and links for 'Home' and 'Databases'. Below this is a search form with a text input field labeled 'Enter Search', a 'SEARCH' button, and a 'List' button. There are also radio buttons for search options: 'All words' (selected), 'Any word', 'Contains text', 'Exact text', and 'Allow operators'. A blue 'Advanced Search' button is located below the search options. Below the search form, there is a message: 'First time user? Need help? Click here for a short guide on using GoGet. If you have any comments or if you are interested in index your radiocarbon data, please Contact us. We greatly appreciate your feedback.' At the bottom, there is a footer with links for 'Help', 'Databases', 'Laboratori d'arqueologia quantitativa (laqu) 2018', and 'Contact Info'.

Figura 4. Interfaz de usuario de Goget

4.2.4. Uso de la aplicación por el usuario

En diciembre de 2019 la base de datos interna de GoGet contenía 193 390 dataciones procedentes de 13 bases de datos internacionales. Más que un rastreador anónimo, GoGet solicita acceso a una base de datos existente para indexarla y permitir las búsquedas. Entre las bases de datos actualmente indexadas debemos citar la Base de datos cronométricos de Catalunya. Creada en 2008, como un esfuerzo conjunto del Museu Nacional d'Arqueologia de Catalunya y el Departamento de Prehistoria de la Universidad Autónoma de Barcelona, contiene registros de las dataciones y contextos datados en yacimientos arqueológicos gestionados por la Generalitat de Catalunya e incluidos en su carta arqueológica. Contiene más de 800 registros procedentes de 157 sitios arqueológicos. Irregularidad en la financiación hace que la base de datos no sea exhaustiva y que no se hayan introducido todavía dataciones posteriores al año 750 ANE. Además de estos registros, GoGet accede a: la base de datos de radiocarbono de la prehistoria de las Islas Baleares (seleccionada y conformada por Rafael Micó); a la base de datos de la prehistoria gallega (seleccionada y conformada por Josefa Rey); a EUBAR, base de datos de la Edad del Bronce en Europa Occidental, seleccionada y conformada por Giacomo Capuzzo, (Capuzzo et al. 2014); a la base de datos de contextos neolíticos de la Península Ibérica, Francia y Suiza compilada por Berta Morell. También indexa los contenidos de varias bases de datos globales: EUROVOL (Manning et al. 2015); aDRAC (dataciones de África central, con datos seleccionados y conformados por Dirk Seidensticker); AustArch, una base de datos australianos, seleccionada y conformada por Alan Williams y Sean Ulm, accesible desde el Archaeology Data Service; la base de datos cronométricos de Gran Bretaña e Irlanda, (también accesible desde el Archaeological Data Service); BANADORA, Base de Datos de radiocarbono francesa y con referencias a sitios europeos y de próximo oriente (<https://www.arar.mom.fr/banadora/>); RADON, la base de datos referida a Europa Central y Escandinavia con dataciones de sitios neolíticos y de la Edad del Bronce (<https://radon.ufg.uni-kiel.de/> mantenida por Martin Hinz y Christoph Rinne); el 14SEA Project, de Agathe Reingruber y Laurens Thissen, <http://www.14sea.org/>), que contiene dataciones del Neolítico y Calcolítico de la Europa Sudoriental y Anatolia. Finalmente, GoGet es capaz de indexar datos de CARD. En la consulta más simple, el usuario introduce un texto, por ejemplo "bone" y obtiene un listado de resultados con la información contenida en la (o las) bases de datos indexadas. Seleccionando cada uno de los resultados obtenidos, GoGet redirecciona al usuario a la base de datos original. Como el sistema busca en numerosos lugares, resulta obvio que en los listados

de resultados también aparecerán muchas repeticiones, que ocurre cuando la misma datación está incluida en diferentes repositorios, a veces con informaciones diferentes o incluso contradictorias (errores). El sistema ofrece al usuario la posibilidad de eliminar automáticamente las repeticiones, seleccionando los datos del repositorio que considere más fiable, o bien deja la posibilidad de que el usuario edite esas repeticiones. Dependiendo de lo genérica que sea la palabra o palabras usadas en la interrogación básica, nos arriesgamos a listados de decenas de miles de dataciones con miles de referencias repetidas

El usuario debe acceder a Goget a través de su navegador a la URL <http://ibercrono.org/goget>. El servidor web le presenta un formulario muy al estilo de Google con una ventana para introducir la información y unos modificadores sobre los términos buscados (figura 5).

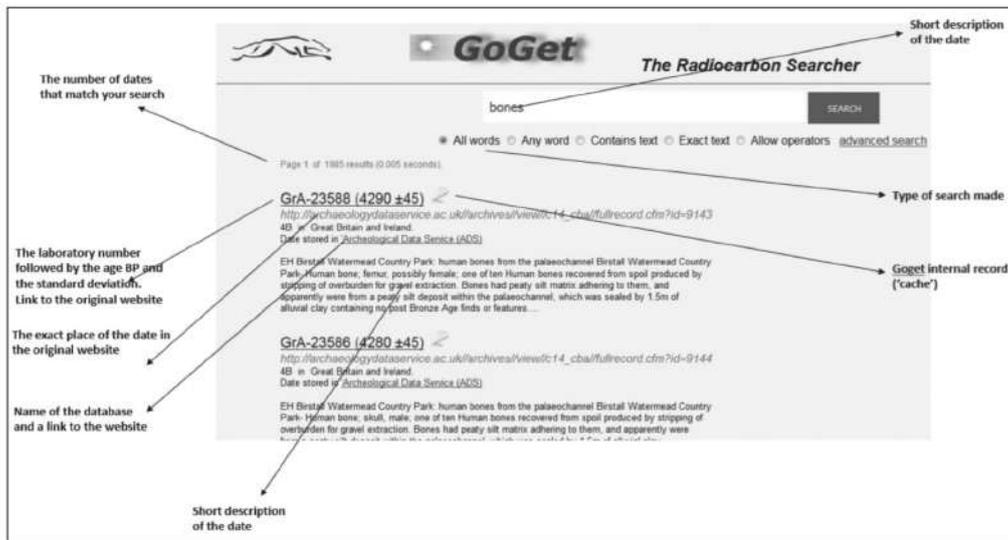


Figura 5. Utilización de Goget

El resultado es una lista de dataciones radiocarbónicas que cumplen el criterio de la información pedida por el usuario. En el ejemplo, el usuario ha buscado por la palabra Barcelona, y le aparece una lista con 179 dataciones radiocarbónicas que continen esa palabra (figura 6).

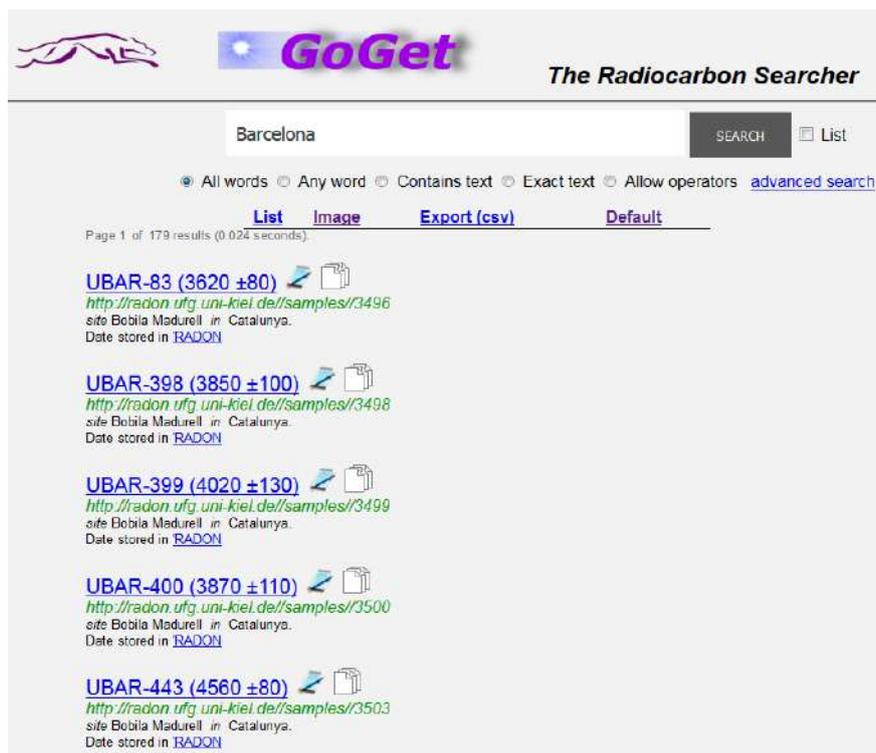


Figura 6. Resultado de una consulta

El usuario sólo debe pulsar sobre la datación para obtener la información completa en la base de datos original. De hecho la aplicación le ‘redirige’ a la ficha dentro del sitio web. Permite consultar los duplicados, si existen e incluso dicha información puede ser descargada. Goget tiene una característica adicional que donde representada de forma gráfica, directamente, el resultado de la búsqueda (figura 7).

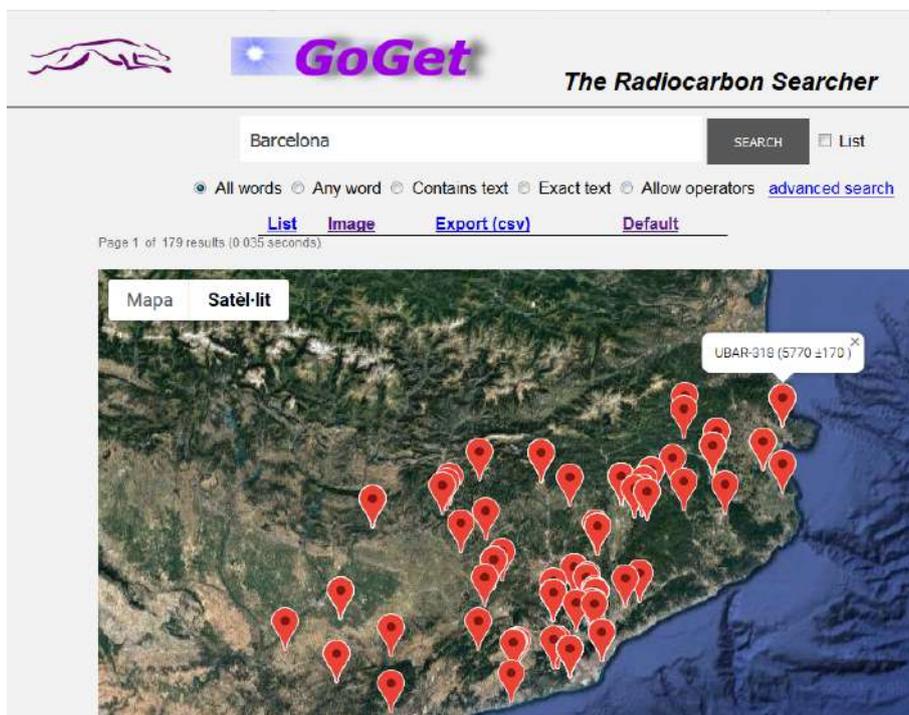


Figura 7. Visualización de una consulta

A menudo la comunidad radiocarbónica necesita de consultas más complejas a la base de datos. Para ello goget cuenta con dos mecanismos complementarios. La búsqueda por operadores y la búsqueda avanzada

- Búsqueda por operadores

La base de datos cuenta con operadores que amplian la búsqueda a partes de texto, conjuntos de palabras o que no contengan ciertas palabras. Así, por ejemplo, la búsqueda de `+Lodge -Farm`, indicando que la consulta se realiza con operadores (figura 8), permite obtener todas las dataciones que tienen la palabra Lodge y que al mismo tiempo no contienen la palabra Farm.

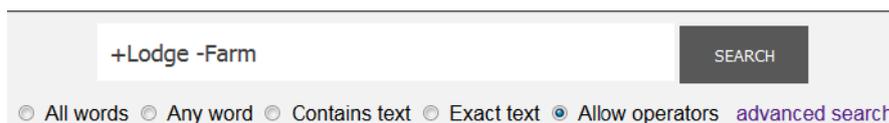


Figura 8. Consulta con operadores

- Búsqueda avanzada

La segunda posibilidad consiste en usar un formulario de búsqueda avanzada empleando un programa asistente inteligente. El asistente inteligente ofrece las siguientes opciones para refinar la búsqueda:

- Base(s) de datos – para limitar la operación de búsqueda en una o más de una de

las bases de datos a las que el sistema tiene acceso. Las bases de datos que se muestran se cargan de manera dinámica a partir de la información contenida en la tabla *bases_datos* en la base de datos MySQL en GoGet. Cada vez que se indexe una nueva base de datos o repositorio, aparecerá de manera automática en la lista.

- Intervalo cronológico (*Time range*).
- Temática. Se puede realizar una búsqueda por tipo de material, tipo de método, o incluso por adscripción cultural o tipo de material característico (p. e. cerámica campaniforme, cerámica campaniana, etc.). Nuevamente, la cuestión del idioma es relevante aquí. GoGet traduce y uniformiza al inglés todos los términos descriptivos, de acuerdo con un diccionario multilingüe accesible y modificable de acuerdo con necesidades específicas.
- Zona Geográfica

Figura 9. Consulta avanzada

5. Goget y las bases de datos cronométricas online

Los campos de la base de datos de Goget pueden agruparse en categorías (figura 10). Una se corresponde con aquella información concierne al origen de los datos indexados. La segunda a la información relativa a la datación radiocarbónica, y la tercera contiene el contexto de la datación. Desafortunadamente esta agrupación no están bien creada en la mayoría de bases de datos online, ya que la información contextual requiere establecer relaciones entre los datos, por lo que su origen no puede ser capturado dentro de un flat file, sino en una base de datos relacional completa.

Website data
Database/website name
Internet address (URL)

Stratigraphy

Sample data
 Sample number
 Laboratory
 Description
 material

Figura 10. Categorías de datos dentro de Goget

El resultado de que actualmente la gran mayoría de las bases de datos de radiocarbono no contienen información sobre el contexto de la datación es que muchas de las inferencias cronológicas no pueden llevarse a cabo. Goget también adolece de este problema.

En esta línea, empiezan a aparecer bases de datos con ontologías completas. A modo de ejemplo, la base de datos de radiocarbono de Catalunya <http://ibercrono.org/cat14> (figura 11) está construida sobre una ontología que contempla estas relaciones (figura 12,13).



Figura 11. Portal web de la base de datos de Radiocarbono de Catalunya

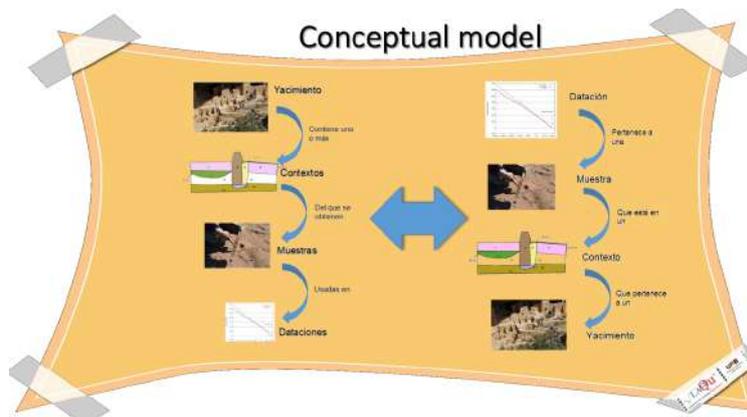


Figura 12. Relaciones entre conceptos ligados a una datación.

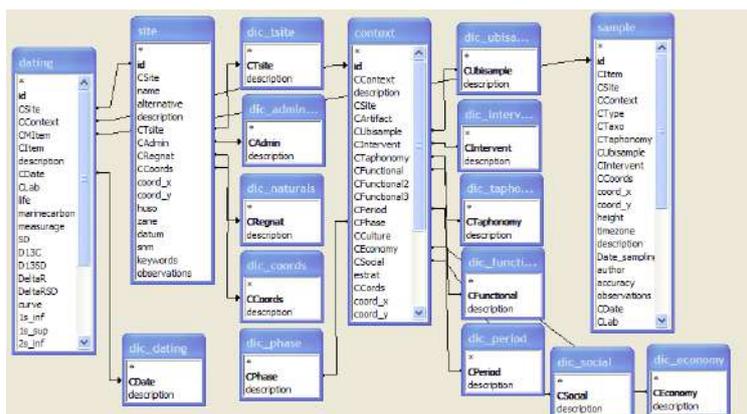


Figura 13. Base de datos de la ontología completa de dataciones y contexto.

6. Futuras mejoras a Goget

El estado actual de desarrollo de goget constituye solo una etapa. Algunas de las mejoras que deben realizarse són muy importantes para un uso pleno de aplicación.

- Incorporar diccionarios de traducción. Las bases de datos pueden estar en distintos idiomas, lo que dificulta la búsqueda de dataciones por palabras. La etapa de carga de una base de datos de Internet a la base de datos debe contener una tarea adicional. Construir un diccionario a fin de homogenizar al menos un diccionario con las palabras más importantes de la información radiocarbónica.
- La bibliografía asociada a una datación, lo que permitiría búsquedas por investigadores, artículos, etc.
- Seguimos con la integración de más bases de datos. Se nos ha revelado como un proceso más lento de lo esperado dado que el *crawler* deber estar confeccionado para la base de datos en concreto.
- Consideramos que los asistentes de consulta basados en IA (Inteligencia Artificial) pueden ayudar a filtrar y obtener de forma más rápida y sencilla la información que el usuario desea.
- La realización de consultas complejas donde se incorpore información temporal, probabilística y geográfica requiere de ampliaciones a los sistemas de consulta actuales. Para ello hay que usar extensiones del sistema de consulta SQL a, por ejemplo, TLSQL2.

Uno de los problemas a los que debe enfrentarse un motor de búsqueda pretendidamente “universal” es la falta en la actualidad de un formato específico para codificar la edad ^{14}C calibrada. No se trata de un escalar (valor numérico), sino de un intervalo de probabilidad. Programas de calibración como OxCal han generalizado un formato tabular para representarlo, si bien esa tabla se divide de manera diferente para diferentes muestras, según sea la probabilidad de los subintervalos dentro del intervalo de calibración definido por 2 veces la desviación típica. En este orden de cosas, la manera más simple de representar el intervalo de confianza de una edad ^{14}C calibrada implica el uso de dos atributos, cada uno definido por un único valor numérico: extremo inferior/extremo superior del intervalo de confianza de la edad calibrada, especificando si se trata del intervalo definido por 1 desviación típica (68% de confianza) o 2 desviaciones típicas (95% de confianza). Es importante utilizar un formato de fecha normalizado, es decir, año calendárico astronómico (con año 0), y usando valores negativos para “años antes de nuestra era” y valores positivos para “años de nuestra era” (cf. capítulos 2 y 11). Cualquier otra forma de codificar las posiciones temporales en la escala calendárica daría problemas a la hora de combinar la información. Del mismo modo, la propia definición de “periodo histórico” plantea problemas de definición, ya que no coincide exactamente con el uso del término “fase” en tanto que conjunto de sucesos arqueológicos, y la mayoría de programas de calibración radiométrica (OxCal, Calib, CalPal, ChronoModel) no lo utilizan. Sin embargo, continúa siendo un campo de información fundamental a la hora de ordenar y filtrar los resultados de una búsqueda de dataciones.

En una próxima versión de GoGet, el campo “periodo histórico” aparecerá en tanto que etiqueta conceptual que defina la duración del suceso arqueológico en términos de un listado específico de intervalos temporales. Dicho listado debe estar referido a un diccionario estable, tipificado, en el que cada entrada está específicamente definida y únicamente codificada. Recientemente se ha producido un interés por definir el término “periodo histórico” de manera formal y precisa. En general, se enfatizan dos maneras de representar rigurosamente estas etiquetas temporales:

- Una referencia semántica no ambigua (etiqueta conceptual) que hace referencia a la vinculación de todos aquellos sucesos que se dice que constituyen el suceso.
- Una referencia temporal no ambigua, que sitúa el inicio y el final del periodo en la escala temporal calendárica.

Y se requieren ambas, ya que son descripciones complementarias, no alternativas. En primer lugar, la *referencia semántica* o etiqueta conceptual –por ejemplo “Chatelperroniense”, “Casapedrense”, “Campaniense”, “Campaniforme”, “Bronce Final IIIb”, “Edad del Hierro”– describe textualmente el periodo exactamente como se indica en la fuente original. Aparentemente consiste en una simple cadena alfanumérica literal, sin etiqueta de idioma, pero debe vincularse al conjunto de referencias bibliográficas concretas que explica la naturaleza de la agregación de sucesos. No es lo mismo la definición de “calcolítico” dada por V. G. Gordon Childe que por E. Sangmeister, o por C. Renfrew, pongamos por caso.

GoGet remite al proyecto *ChronOntology* (<https://chronontology.dainst.org/>), con sede en el Instituto Arqueológico Alemán (DAI) y que forma parte de la estructura iDAI.welt, el cual tiene como propósito fundamental crear un servicio web abierto para el direccionamiento dinámico de toda la gama de conceptos cronológicos, permitiendo asignar referencias temporales estandarizadas a los recursos digitales de humanidades digitales y los vincula con otros sistemas mundiales de nomenclátor temporal (Thiery y Engel 2016, Thiery y Mees 2016a, 2016b, 2018). Proporciona una herramienta para recopilar y vincular referencias a “periodos históricos” que se encuentran en publicaciones u otras fuentes. No busca definir periodos o asignarles fechas si la información no se encuentra en las fuentes originales, ni tampoco busca resolver problemas relacionados con el conocimiento o tomar decisiones sobre lo que es una buena metodología de conocimiento. Un aspecto importante del modelo de datos de ChronOntology es la necesidad de especificar el tipo de “periodo” al que hacemos referencia. Cada periodo representa un intervalo de tiempo específico marcado por características, eventos o cambios específicos en algún lugar del planeta. Por ejemplo, una era en geología denota un periodo de tiempo claramente definido de duración bien definida, delimitado por un evento de inicio y un evento final; en la historia social, las eras pueden denotar el intervalo de tiempo durante el cual un monarca ejerció su reinado; en las artes, un periodo es un periodo de tiempo en el que al principio y al final las prácticas (o modas) cambian en un grado significativo. Por lo tanto, lo que define un periodo puede ser muy variado, y es posible que haya periodos superpuestos que ocupen el mismo volumen de espacio-tiempo pero que estén definidos por diferentes disciplinas. “Tipo” es el campo en el que se especifica la característica definitoria del periodo en términos sencillos. Lo ideal sería que solo se definiese una clase, pero se puede seleccionar más de una en los casos en que no sea posible identificar una clase de periodo. En caso de que la característica definitoria del periodo no esté clara en la fuente original (es decir, no se indica y no puede deducirse), puede utilizarse el tipo “no especificado”. Una ontología alternativa a *ChronOntology* como *PeriodO* (<http://perio.do/en/>) (Rabinowitz et al. 2016, 2018) exige la formalización del intervalo temporal de cada periodo u época de acuerdo con el estándar *OWL-Time* (<https://www.w3.org/TR/owl-time/>). Según esta normativa, una definición de periodo debe incluir necesariamente un intervalo de tiempo con sus puntos iniciales y finales bien definidos. Se utiliza la etiqueta `time:year` para descripciones de intervalos que pueden representarse con un solo año (el año “600 A.C.” se describiría con un valor de `hora:año` de -0599), o bien, la etiqueta `period:earliestYear` y `period:latestYear` para descripciones de intervalos que necesitan ser representados como rangos. Por ejemplo, un intervalo con la descripción textual “ocho siglos antes de Cristo” puede describirse con una descripción de la fecha que tenga un `period:valor` del año más antiguo de -0799 y un `period:valor` del año más reciente de -0700. El sistema también acepta periodos con especificaciones de tiempo o espacio que faltan o están incompletas para permitir un reflejo preciso del estado de la investigación. Por ejemplo, un periodo puede no estar bien determinado en el tiempo y/o el espacio, o tal información no se menciona explícitamente en las fuentes. La etiqueta `time:intervalStartedBy` enlaza la definición del periodo con un intervalo de tiempo que tiene el mismo punto de inicio que el periodo y un punto final (desconocido) que precede al punto final del periodo. Este sería el intervalo de inicio para el periodo. Por su parte, la etiqueta `time:intervalFinishedBy` enlaza la definición del periodo con un intervalo de tiempo (anónimo)

que tiene el mismo punto final (desconocido) que el periodo y un punto inicial (desconocido) que se encuentra después del punto inicial del periodo. Constituiría el intervalo de finalización (*stop*) para el periodo. Goget implementará próximamente un diccionario de “periodos históricos” para cada una de las bases de datos en las que busca dataciones, formalizando su expresión de acuerdo con los formatos propios de *ChronOntology* y *PeriodO*

Bibliografía

Allen, James F. "[Maintaining knowledge about temporal intervals](#)". *Communications of the ACM*. **26** (11): 832–843 [doi:10.1145/182.358434](#). [ISSN 0001-0782](#).

Antova L., Jansen T., Koch C. and Olteanu D. "Fast and Simple Relational Processing of Uncertain Data". *Proc. 24th International Conference on Data Engineering, ICDE 2008, April 7-12, 2008, Cancun, Mexico*, pp. 983-992.

Barceló J.A., Bogdanović, I., Vicente, O., 2004, A Theory of Knowledge Building by Using Internet. In K. Fennema and H. Kamermans (Eds.), *Making the Connection to the Past*. Proceedings of the 27th Computer Applications and Quantitative Methods in Archaeology (CAA) Conference held in Dublin in 1999. Leiden: Leiden University Press, (pp. 31-36).

Barceló, J.A., Bogdanović, I., Piqué, R., 2004., *Telearchaeology*, *Archeologia e Calcolatori*, 15, Roma, pp. 467-481.

Bogdanović, I., 2002, *Internet, el conocimiento distribuido y la resolución del problema arqueológico*, Master thesis - unpublished;

Bogdanovic, I, Capuzzo, G., Barceló, J.A. 2013, A database for radiocarbon dates. Some methodological and theoretical issues about its implementation. En *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* Editado por Earl, Graeme, Tim Sly, Angeliki Chrysanthi, Patricia Murrieta-Flores, Constantinos Papadopoulos, Iza Romanowska and David Wheatley. [Amsterdam University Press, Amsterdam](#) (ISBN 9789089646637).

Codd E. F.. 1970. A relational model of data for large shared data banks. *Commun. ACM* 13, 6 (June 1970), 377-387. DOI: <https://doi.org/10.1145/362384.362685>

Cox, Simon & Little, Chris. (2016). *Time Ontology in Owl*.

Dan Suciú, Dan Olteanu, R. Christopher, and Christoph Koch. 2011. *Probabilistic Databases* (1st ed.). Morgan & Claypool Publishers.

Flesca Sergio, Furfaro Filippo, Parisi Francesco. 2014, Consistency checking and querying in probabilistic databases under integrity constraints. *Journal of Computer and System Sciences*, Volume 80, Issue 7, Pages 1448-1489.

Fowler, Glenn S. "cql - A Flat File Database Query Language." *USENIX Winter* (1994).

Hansen, Thomas & Cordua, Knud & Zunino, Andrea & Mosegaard, Klaus. (2016). *Probabilistic Integration of Geo-Information*. 10.1002/9781118929063.ch6.

Holdaway, Simon J. and Wandsnider, LuAnn, "Time in Archaeology: An Introduction" (2008). *Anthropology Faculty Publications*. 81. <http://digitalcommons.unl.edu/anthropologyfacpub/81> (consultado 25 de septiembre de 2019)

Hong Zhu, Caicai Zhang, Zhongsheng Cao, and Ruiming Tang. 2016. On efficient conditioning of probabilistic relational databases. *Know.-Based Syst.* 92, C (January 2016), 112-126. DOI=<http://dx.doi.org/10.1016/j.knosys.2015.10.017>

Jensen C.S., Soo M.D. and Snodgrass R. T. Unifying Temporal Data Models via a Conceptual Model. *Inf. Syst.*, 19(7):513–547, December 1994.

Jull, A., & McNichol, A. (2014). Editorial: Conventions for Reporting Radiocarbon Determinations. *Radiocarbon*, 56(2), Ix-X. doi:10.1017/S003382220004941

Kra, R. (1988). Updating the past: The Establishment of the International Radiocarbon Data Base. *American Antiquity*, 53(1), 118-125. doi:10.2307/281158

Kulkarni, Krishna, and Jan-Eike Michels. "[Temporal features in SQL: 2011](#)". ACM SIGMOD Record 41.3 (2012): 34-43.

Levene ark. 2010. *An Introduction to Search Engines and Web Navigation* (2nd ed.). Wiley Publishing.

Levine, Abigail & Stanish, Charles. (2013). The Importance of Multiple 14C Dates from Significant Archaeological Contexts. *Journal of Archaeological Method and Theory*. 21. <https://doi.org/10.1007/s10816-013-9177-4>

Millard, A. (2014). Conventions for Reporting Radiocarbon Determinations. *Radiocarbon*, 56(2), 555-559. doi:10.2458/56.17455

Morris, G., & Conner, L. M. (2017). Assessment of accuracy, fix success rate, and use of estimated horizontal position error (EHPE) to filter inaccurate data collected by a common commercially available GPS logger. *PloS one*, 12(11), e0189020.

Nagori, Viral & Trivedi, Bhushan (April 2014). Types of Expert System: Comparative. *Asian Journal of Computer and Information Systems* (ISSN: 2321–5658) Volume 02–Issue 02

Nilesh Dalvi, Christopher Re, and Dan Suciu. 2011. Queries and materialized views on probabilistic databases. *J. Comput. Syst. Sci.* 77, 3 (May 2011), 473-490. DOI=<http://dx.doi.org/10.1016/j.jcss.2010.04.006>

Papaioannou, Katerina & Theobald, Martin & Hanspeter Böhlen, Michael. (2018). Supporting Set Operations in Temporal-Probabilistic Databases. 1180-1191. 10.1109/ICDE.2018.00109.

Parisi, Francesco and Grant, John. 2017. On repairing and querying inconsistent probabilistic spatio-temporal databases. *Int. J. Approx. Reasoning* 84, C (May 2017), 41-74. DOI: <https://doi.org/10.1016/j.ijar.2017.02.003>

Revesz, Peter. (2010). Introduction to databases. From biological to spatio-temporal. 10.1007/978-1-84996-095-3.

R.L. Ackoff, From data to wisdom, *Journal of Applied Systems Analysis* 16 (1989) 3–9.

Rowley, J. (2007). The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of Information Science*, 33(2), 163–180. <https://doi.org/10.1177/0165551506070706>

Shawn Graham, Neha Gupta, Jolene Smith, Andreas Angourakis, Andrew Reinhard, Kate Ellenberger, Zack Batist, Joel Rivard, Ben Marwick, Michael Carter, Beth Compton, Rob Blades, Cristina Wood, & Gary Nobles (2019). The Open Digital Archaeology Textbook. Consultable via internet <https://o-date.github.io/draft/book/index.html> (consultado el 25 de septiembre 2019)

Snodgrass, R.T., editor, *The TSQL2 Temporal Query Language*, Kluwer Academic Publishers, 1995, 674+xxiv pages.

Snodgrass R.T. *Developing Time-Oriented Database Applications in SQL*. Morgan Kaufmann, San Francisco, CA, July 1999.

Stravoskoufos, Konstantinos & Petrakis, Euripides & Mainas, Nikolaos & Batsakis, Sotirios & Samoladas, Vasilis. (2016). SOWL QL: Querying Spatio-Temporal Ontologies in OWL. *Journal on Data Semantics*. 5. 10.1007/s13740-016-0064-5.

T. R. Gruber. A translation approach to portable ontologies. *Knowledge Acquisition*, 5(2):199-220, 1993.

Wood, Rachel. (2015). From revolution to convention: The past, present and future of radiocarbon dating. *Journal of Archaeological Science*. 56. 10.1016/j.jas.2015.02.019.

Zaniolo, Carlo. (2009). Event-Oriented Data Models and Temporal Queries in Transaction-Time Databases. *TIME 2009 - 16th International Symposium on Temporal Representation and Reasoning*. 47-53. 10.1109/TIME.2009.26.