



Preguntas frecuentes sobre el Archivo de la Web Española

1. ¿Cómo puedo acceder a los sitios web archivados?
2. ¿Por qué existen restricciones de acceso a sitios web que están libremente disponibles a través de Internet?
3. ¿Puedo descargar o imprimir las páginas web archivadas?
4. ¿Cómo funciona la herramienta de consulta del Archivo Web?
5. ¿Se pueden realizar búsquedas por palabras clave?
6. ¿Qué otra información puedo obtener sobre la recolección de un sitio web?
7. ¿Por qué no encuentro lo que busco?
8. ¿Los sitios web se visualizan tal y como aparecían el día de la captura?
9. ¿Por qué no consigo visualizar determinados elementos de las páginas web (imágenes, galerías, vídeos, etc.)?



1. ¿Cómo puedo acceder a los sitios web archivados?

Para acceder a los sitios web archivados debe dirigirse a la Biblioteca Nacional de España o a las bibliotecas habilitadas por las Comunidades Autónomas para ello por ser depositarias del depósito legal. En ellas se han habilitado ordenadores específicos para la consulta de esta colección.

Sin embargo, desde internet o desde terminales no autorizados, es posible acceder a la herramienta de búsqueda (<http://www.dl-e.es/openwayback/wayback/>) y ver si un sitio web ha sido archivado y en qué fechas se ha capturado.

2. ¿Por qué existen restricciones de acceso a sitios web que están libremente disponibles a través de Internet?

Por respeto a la legislación en materia de propiedad intelectual. El autor que ha creado los contenidos archivados tiene el derecho de decidir dónde y cómo darles difusión. Para garantizar este derecho, los centros de conservación están obligados a limitar la consulta de los sitios web archivados a los terminales habilitados. [Véase [¿Cómo puedo acceder a los sitios web archivados?](#)]

3. ¿Puedo descargar o imprimir las páginas web archivadas?

El acceso a los sitios web archivados debe respetar la legislación vigente sobre propiedad intelectual, por lo que no se permiten descargas, copias o transferencia de ficheros por vía directa. Por eso el acceso a esta colección debe realizarse a través de terminales específicos que impidan estas prácticas ilegales.

Próximamente se habilitarán los procedimientos para la obtención de reproducciones, que en todo caso podrán hacerse exclusivamente con fines de investigación.

4. ¿Cómo funciona la herramienta de consulta del Archivo Web?

Desde un terminal autorizado [véase [¿Cómo puedo acceder a los sitios web archivados?](#)], es posible explorar todos los sitios web archivados. Para hacerlo es necesario conocer la dirección o URL del recurso que se pretende consultar e introducirla en la barra de búsqueda.

Después de hacer clic en el botón *¡Llévame al pasado!* se accede al calendario de las recolecciones. El calendario muestra el número de veces y los días en los que el recurso solicitado ha sido capturado, entre 2009 y la actualidad. Basta con hacer clic sobre uno de los días para visualizar la página web con los elementos que contenía en el momento de la captura. El documento que se visualiza no es una mera imagen estática, sino una reproducción de la página web que mantiene sus funciones de navegación y puede contener texto, imágenes, vídeos, hiperenlaces, etc.

La captura intenta proporcionar una experiencia de navegación lo más parecida posible a la original, tanto por sus contenidos como por su diseño y apariencia. Sin embargo, en algunos casos, no se archiva la totalidad de los contenidos de un sitio web, ni se pueden evitar los errores de apariencia de las páginas [véase [¿Los sitios web se visualizan tal y como aparecían el día de la captura?](#)]. Se trata de una tecnología y de un servicio en fase de pruebas.

5. ¿Se pueden realizar búsquedas por palabras clave?



De momento, la única forma de buscar un recurso es por URL, tanto en la búsqueda simple como en la búsqueda avanzada. Próximamente se facilitará la posibilidad de buscar por materia y por título, así como a texto completo por palabra clave y por colección.

6. ¿Qué otra información puedo obtener sobre la recolección de un sitio web?

Las características principales de un sitio web archivado son la frecuencia (periodicidad de la recolección: diaria, semanal, mensual, etc.) y la profundidad (el número de clics necesarios para alcanzar los recursos de un sitio web a partir de la URL principal). Es posible consultar la frecuencia con la que un recurso ha sido archivado en el calendario que aparece tras una búsqueda en la interfaz de búsqueda. Haciendo clic en uno de los días marcados en azul se accede al recurso solicitado.

En cambio, para verificar la profundidad de los sitios web se deben consultar los ficheros Excel con las semillas incluidas en cada colección, disponibles en la página web de la colección del Archivo de la Web Española

7. ¿Por qué no encuentro lo que busco?

Puede deberse principalmente a dos razones. En primer lugar, la URL que se ha introducido podría ser incorrecta o estar incompleta: de momento, no es posible realizar búsquedas a texto completo. En segundo lugar, es posible que el sitio web o parte del sitio no hayan sido recolectados por incidencias de carácter técnico, o por no responder a los criterios de selección elegidos.

Hay que tener en cuenta que el Archivo de la Web Española conserva miles de sitios web, pero sus colecciones no pueden ser exhaustivas y varían según las características de recolección [Véase [¿Qué otra información puedo obtener sobre la recolección de un sitio web?](#)].

Por lo general, la Biblioteca Nacional de España asegura por lo menos una captura anual y automática de todos los sitios web con dominio .es: esta “recolección masiva” pretende guardar una muestra representativa que refleje el panorama de Internet en un momento dado. Sin embargo, no se guardan todos los niveles de un sitio web, así que algunos hiperenlaces podrían no funcionar.

Asimismo, la Biblioteca Nacional de España y los centros de conservación de las Comunidades Autónomas realizan la captura del patrimonio documental relacionado con España, también para los dominios distintos al .es. Las llamadas “recolecciones selectivas” son colecciones específicas con recursos seleccionados sobre un tema o sobre un acontecimiento de interés. Recogen menos sitios web, pero con una frecuencia y una profundidad mayores que en el caso de las masivas. Aun así, queda la posibilidad de no encontrar una página web o uno de sus elementos: porque no ha sido incluido en la recolección selectiva; porque no se ha realizado su captura el día deseado; o porque, por razones técnicas o de espacio, el recurso solicitado no está archivado.

8. ¿Los sitios web se visualizan tal y como aparecían el día de la captura?

El propósito de las recolecciones web no es sólo conservar los contenidos sino también la apariencia y la interactividad de cada página. Sin embargo, en las recolecciones pueden faltar datos o darse otros inconvenientes: algunos hiperenlaces podrían redirigir a recursos no archivados; otros objetos (imágenes, vídeos, pdf, etc.) podrían presentar errores de visualización; los elementos que necesitan de interacción humana (filtros, búsquedas en el sitio, etc.) no funcionan; etc.

Además, la página web podría presentar errores de visualización, debido al diseño de algunas páginas web o al uso de un lenguaje de programación que no es posible reproducir. Por ello, la página archivada podría no respetar la configuración y la apariencia del recurso original. En ocasiones, los defectos de visualización pueden afectar a la navegación; pero, en la mayoría de los



casos, no impiden la consulta de los contenidos de interés. La solución de todos estos problemas sigue siendo objeto de investigación y desarrollo en el campo del archivado web a nivel internacional. Finalmente, hay que tener en cuenta que la profundidad de recolección (el número de clics necesarios para alcanzar los recursos de un sitio web a partir de la URL principal) no es ilimitada, está condicionada por la configuración establecida dentro de los objetivos de recolección de cada colección en particular.

9. ¿Por qué no consigo visualizar determinados elementos de las páginas web (imágenes, galerías, vídeos, etc.)?

La correcta visualización de los contenidos de una página web depende principalmente de dos factores: por un lado, de la efectividad en la captura del sitio; por otro lado de la capacidad de la herramienta de consulta para reproducir los contenidos recolectados. En este segundo caso, aunque los datos estén archivados, puede ocurrir que los datos se hayan archivado, pero que su visualización resulte, de momento, imposible por las limitaciones técnicas del software de recuperación. Los elementos que presentan este tipo de dificultades para su visualización son principalmente aquellos que utilizan javascript y los vídeos en streaming.