

Soluciones software para sistemas de información web en humanidades digitales: revisión, análisis y estudio comparativo

Software solutions for web information systems in digital humanities: review, analysis and comparative study

Maurizio Toscano; Manuel J. Cobo; Enrique Herrera-Viedma

Note: This article can be read in its English original version on:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/86613>

Cómo citar este artículo.

Este artículo es una traducción. Por favor cite el original inglés:

Toscano, Maurizio; Cobo, Manuel J.; Herrera-Viedma, Enrique (2022). "Software solutions for web information systems in digital humanities: review, analysis and comparative study". *Profesional de la información*, v. 31, n. 2, e310211.

<https://doi.org/10.3145/epi.2022.mar.11>

Artículo recibido el 18-07-2021
Aceptación definitiva: 14-02-2022



Maurizio Toscano ✉
<https://orcid.org/0000-0001-5418-3859>

Fecyt
Oficina Europea
C/ Pintor Murillo, 15
28100 Alcobendas (Madrid), España
maurizio.toscano@fecyt.es



Manuel J. Cobo
<https://orcid.org/0000-0001-6575-803X>

Universidad de Granada
Departamento de Ciencias de la
Computación e Inteligencia Artificial
Instituto de Investigación DaSCI
Periodista Daniel Saucedo Aranda, s/n
18014 Granada, España
mjcobo@ugr.es



Enrique Herrera-Viedma
<https://orcid.org/0000-0002-7922-4984>

Universidad de Granada
Departamento de Ciencias de la
Computación e Inteligencia Artificial
Instituto de Investigación DaSCI
Periodista Daniel Saucedo Aranda, s/n
18014 Granada, España
viedma@decsai.ugr.es

Resumen

La investigación en humanidades depende cada vez más de cómo se estructura y gestiona la información y de cómo, a partir de ella, se produce nuevo conocimiento. Además, los enfoques participativos, que a menudo utilizan los sistemas de información de la web como su infraestructura de soporte, han influido en las tendencias historiográficas más recientes, en particular en el marco metodológico de las humanidades digitales. El objetivo de este trabajo ha sido realizar, desde una perspectiva operativa y de implementación, una revisión de soluciones software comúnmente utilizadas para el desarrollo de sistemas de información web en proyectos de investigación en humanidades y patrimonio cultural, con el fin de proporcionar una visión de las distintas opciones disponibles, con sus aspectos positivos y sus limitaciones, también en función de diferentes necesidades a nivel usuario. Se ha llevado a cabo un análisis individual y comparativo de dieciséis paquetes software comúnmente utilizados en estos campos, ya sean genéricos o desarrollados para un dominio de investigación específico, considerando sus principales funciones, fortalezas y debilidades. Los resultados obtenidos facilitan una toma de decisiones crítica y razonada entre varias opciones disponibles, ofreciendo orientación a los creadores de esos sistemas, tanto investigadores como desarrolladores, y proporcionando también una base común de términos y casos de uso para facilitar su necesario diálogo.

Palabras clave

Sistemas de información web; Humanidades digitales; Sistemas de gestión de contenidos; Sistemas de gestión de activos digitales; Entornos virtuales de investigación; Crowdsourcing; Análisis comparativo; Revisión de software.

Abstract

Research in the humanities increasingly depends on how information is structured and managed and how, on the basis of that information, new knowledge is produced. Additionally, participatory approaches, which often rely on web information systems as their supportive infrastructure, have made an impact on the most recent historiographical trends, in particular in the methodological framework of digital humanities. The aim of this paper was to produce, from an operational and implementation perspective, a review of software solutions frequently used to develop web information systems for research projects in humanities and cultural heritage, in order to provide an understanding of the various possibilities available and their positives and limitations, also based on different users' requirements. An individual and comparative analysis of sixteen different application frameworks commonly used in these fields, either generic or developed for a specific research domain, has been carried out, considering their main functionalities, strengths, and weaknesses. The achieved results facilitate critical and reasoned decision-making among several available options, guiding the makers of those systems, both researcher(s) and developers(s), and providing them also with a common ground of terms and use cases to facilitate their necessary dialogue.

Keywords

Web information systems; Digital humanities; Content management systems; Digital assets management; Virtual research environments; Crowdsourcing; Comparative analysis; Software review.

Financiación

Este artículo es un resultado de los proyectos: "Sistemas de toma de decisiones en grupo disruptivos en ambiente difuso: Aplicaciones en gestión inteligente de energía y empleados", 2020-2023, Referencia: PID2019-103880RB-I00, financiado por MCIN/AEI/10.13039/501100011033, IP Enrique Herrera Viedma; y "Nuevos sistemas difusos para la toma de decisiones: Aplicaciones en entornos digitales", 2021-2023, Referencia: P20_00673, financiado por la Junta de Andalucía, IP Enrique Herrera Viedma.

1. Introducción

Un sistema de información es un conjunto de personas, datos, procesos y tecnología que interactúan para recoger, procesar, almacenar y proveer la información necesaria para el correcto funcionamiento de una organización (Whitten *et al.*, 2004). Nacen en el ámbito empresarial para gestionar, almacenar, procesar y recuperar los datos que lo componen, con el fin de producir información. Son entornos colaborativos por definición, que abogan hacia la puesta en común y el trabajo en grupo, cada vez más perentorio en humanidades. Son anteriores a Internet (Hirschheim; Klein, 2012), pero hoy la Web constituye el entorno ideal para desplegar su componente tecnológico, gracias a interfaces de navegación fáciles de usar, que emplean herramientas con las que los usuarios están ampliamente familiarizados, y a su capacidad de operar de forma prácticamente independiente, ya sea a nivel hardware que software, de los sistemas operativos lado servidor y cliente. Además de eso, la Web resulta finalmente un componente habilitador, que sustenta y favorece la apertura de esos sistemas al exterior, hacia otros grupos de investigación o hacia la sociedad (Toscano, 2018).

El campo de aplicación de este estudio son las humanidades digitales (HD), una tendencia relativamente reciente en el estudio del pasado, que destaca a nivel nacional e internacional, ya sea como objeto de investigación o como temática preferente de financiación (Burdick *et al.*, 2013; Toscano *et al.*, 2020). En este contexto, los sistemas de información web (SIW) son una pieza central en el proceso de digitalización de la investigación: a través de las diferentes fases de planteamiento, modelado e implementación, propician, y en cierto modo obligan, el planteamiento de nuevas preguntas y la reformulación de aquellas ya planteadas, lo que en definitiva constituye la cuestión nuclear de las HD en el mundo actual (Rodríguez-Ortega, 2018). Estos sistemas son en muchos casos el punto de partida de ese recorrido y la primera ocasión en la que el investigador que utiliza métodos tradicionales se enfrenta con una serie de problemáticas y oportunidades planteadas por el cambio digital.

En el proceso de desarrollo de un SIW (figura 1), una vez realizado el diseño del modelo de datos y evaluadas las funciones necesarias para cubrir las exigencias específicas, es necesario elegir el entorno de trabajo más apropiado para la sucesiva implementación a nivel informático. Este artículo se centra concretamente sobre este aspecto, a través de una revisión crítica de las soluciones software existentes para la implementación de sistemas de información web en humanidades y gestión del patrimonio cultural. Se procederá con un análisis individual y comparativo de paquetes que suelen utilizarse en esas áreas, sean genéricos o desarrollados para un dominio específico de

Las disciplinas humanísticas ya no pueden prescindir de soluciones TIC para gestionar las crecientes cantidades de datos que generan, y para respaldar la generación de conocimiento a través de enfoques cuantitativos

investigación, y teniendo en cuenta sus principales funciones, ventajas e inconvenientes.

Las disciplinas humanísticas ya no pueden prescindir de soluciones informáticas para gestionar las crecientes cantidades de datos e información que generan, y para respaldar las prácticas asociadas a la generación de conocimiento a través de enfoques cuantitativos, que han crecido progresivamente en popularidad. Cada vez más, la investigación en humanidades depende de cómo se configura, se estructura y se gestiona la información y de cómo, a partir de ella, se produce nuevo conocimiento mediante procesos interpretativos. Además, los enfoques participativos, que con frecuencia se sirven de los SIW como infraestructura informática de apoyo, han hecho mella en las tendencias actuales de la historiografía, en particular en el marco metodológico de la Historia pública y de la Historia pública digital (Gallini; Noiret, 2011; Noiret, 2018; Pons, 2018). A la vez, se ha detectado un creciente interés de la sociedad misma por participar en proyectos co-creativos y colaborativos en humanidades y patrimonio cultural (Bocanegra-Barbecho *et al.*, 2017; Bocanegra-Barbecho, 2020), a través de las redes sociales y en el marco de iniciativas de ciencia ciudadana (Hedges; Dunn, 2018; Ridge, 2014; Terras, 2016). En definitiva, los nuevos historiadores públicos, y los humanistas digitales en general, representan el principal público objetivo de este trabajo, en cuanto promotores y hacedores de recursos e infraestructuras digitales.

Los datos, en particular en las disciplinas que estudian el pasado, no suelen perder actualidad y una vez almacenados de manera estructurada pueden llegar a explotarse y reelaborarse a lo largo de décadas. Al contrario que ocurre con los SIW, que son hijos de sus tiempos y de las tecnologías disponibles y empleadas en un determinado momento, que son contingentes y sufren una rápida evolución, también en función de los dispositivos que se utilizan para interactuar con ellos. Cualquier análisis que se haga al respecto es una instantánea del estado actual de la evolución de la tecnología en ese sector. A pesar de ello, o justamente por ese motivo, es necesario revisar y actualizar el panorama de las soluciones software actualmente disponibles en el mercado y más frecuentemente utilizadas en humanidades o para la gestión de recursos materiales e inmateriales del patrimonio cultural. Basarse sólo en la familiaridad y la experiencia previa con un software específico, que de todos modos son factores subjetivos que merecen ser tenidos en cuenta, conlleva el riesgo de caer en la noción conocida como *golden hammer* (Brown *et al.*, 1998), es decir la aplicación obsesiva de una tecnología familiar para resolver un número demasiado elevado de problemas software. Como afirma el desarrollador de software Gilgado (2014):

“El problema de usar siempre las mismas aplicaciones es que no tienes suficientes argumentos para elegir porque no tienes con qué comparar y te limita el conocimiento”.

El objetivo de este artículo ha sido elaborar una revisión de este tipo de softwares desde una perspectiva operativa y de implementación, con el fin de proporcionar a los lectores una comprensión de las diversas posibilidades disponibles para construir sistemas de información, sus aspectos positivos y sus limitaciones, también en función de los requisitos de los usuarios. El objetivo final es facilitar una toma de decisiones crítica y razonada entre las opciones disponibles, orientando a los creadores de esos sistemas, tanto investigadores como desarrolladores, y proporcionándoles también una base común de términos y casos de uso para facilitar su necesario diálogo (Scheuermann; Kroeze, 2017).

El texto está organizado en cinco capítulos, más allá de esta introducción: la metodología, que introduce los tipos de productos analizados y explica los criterios seguidos para su selección; la descripción de los paquetes software seleccionados, organizada en cuatro secciones; el análisis comparativo, que presenta la discusión de los resultados; un capítulo dedicado a las limitaciones, y finalmente las conclusiones.

2. Metodología

Entre las soluciones software utilizadas para implementar un sistema de información web (SIW), se pueden identificar al menos tres categorías principales:



Figura 2. Infografía de las fases nucleares en la elaboración de un sistema de información web personalizado.

- *Content Management Systems* (CMS);
- *Digital Assets Management* (DAM); y
- *Virtual Research Environment* (VRE).

Los CMS se utilizan principalmente para recopilar, almacenar y publicar en la web contenidos digitales alfanuméricos generados de forma colaborativa por los usuarios, según un determinado flujo de trabajo y en un entorno amigable (Martínez-Caro *et al.*, 2018). Los DAM son un tipo de CMS centrado en la gestión de colecciones de objetos: recursos y medios, junto con sus metadatos, en lugar de contenidos alfanuméricos (Jiménez, 2003). Los VRE son similares a los CMS, pero contruidos específicamente para la investigación, como laboratorios virtuales para respaldar la colaboración y proporcionan instrumentos de análisis y visualización incorporadas (Candela *et al.*, 2013). Debido a sus especificidades y a su relevancia para nuestro dominio de aplicación, decidimos tratar las soluciones software orientadas al *crowdsourcing* (Oomen; Aroyo, 2011) como una categoría adicional independiente.

Con el fin de ofrecer una panorámica suficientemente amplia y coherente de los softwares disponibles en el mercado, hemos establecido unos criterios básicos de selección:

- software gratuito y de código abierto (*Free and Open Source Software* - FOSS), que permite modificaciones, integraciones y adaptaciones personalizadas sin ninguna restricción, para satisfacer las necesidades de los proyectos de DH;
- programas en desarrollo y mantenimiento activo¹;
- con evidencia de casos de uso y aplicaciones demostradas en el entorno de las humanidades o del patrimonio cultural².

Se han excluido de esta revisión, a pesar de ser soluciones muy populares, paquetes software que ofrecen por defecto una muy limitada personalización y donde la gestión del contenido desestructurado prima sobre la gestión y la manipulación de datos y metadatos, como es el caso de *Joomla* o *WordPress*.

Los criterios utilizados para el análisis proceden del examen comparativo de la documentación oficial, combinado con una estrategia adoptada en revisiones similares publicadas recientemente (Nishanbaev *et al.*, 2019). El conjunto resultante es bastante amplio, ya que comprende 15 parámetros, pero eso no excluye que puedan haberse dejado de lado otros factores relevantes.

Se han agrupado en tres bloques:

- 1) El primero incluye las características esenciales, como los formatos de archivo soportados, el mapeo, la interfaz traducible, la importación y exportación masiva, y la posibilidad de ampliar las características a través de módulos o *plugins*.
- 2) El segundo bloque se centra en las funciones de modelado de datos y recuperación de la información, que es el conjunto de características distintivas proporcionado por un SIW (Lenz; Oberweis, 1998).
- 3) El tercer bloque comprende algunas características adicionales que hemos considerado merecedoras de ser analizadas, ya sea porque responden a una interesante tendencia hacia aplicaciones desacopladas (API REST y aplicación móvil) o porque facilitan la adopción inicial y el mantenimiento a largo plazo (respectivamente, versión demo online y *SaaS*³).

Los criterios identificados no están exentos de posibles limitaciones relacionadas, por ejemplo, con el amplio conjunto de programas informáticos seleccionados. Esto significa que no podemos excluir que el análisis, por ejemplo, de cada categoría de software con un conjunto específico de criterios hubiera proporcionado alguna información adicional. En este sentido, hemos decidido favorecer la intercomparabilidad entre un conjunto de soluciones que se utilizan a menudo como alternativas en contextos similares. Decidimos proceder así también porque, gracias a la progresiva incorporación de nuevas funciones, las diferencias entre paquetes de distintas categorías van disminuyendo.

Para llevar a cabo el análisis comparativo se ha adoptado un enfoque combinado para probar cada producto software. En todos los casos en que ha sido posible, el protocolo ha consistido en el análisis de la documentación oficial combinada con el uso de la versión de prueba online. En los pocos casos en los que esta última no se encuentra disponible (tabla 7), el software se ha instalado en un servidor remoto. Además, la mayoría de los programas informáticos seleccionados ha sido ya objeto de revisiones individuales publicadas en artículos de investigación, que también se han tenido en cuenta durante el análisis y se citan en el siguiente capítulo 3.

3. Descripción de los paquetes software seleccionados

Existe una cantidad relevante de soluciones tecnológicas elaboradas y mantenidas para dar respuesta a la necesidad de gestionar información estructurada y documentación en la Web. Mientras que algunas son claramente alternativas entre sí, otras cubren diferentes aspectos, con funciones que se solapan sólo en parte, y son más difícilmente comparables. Un listado representativo, aunque incompleto y temporal, de aplicaciones FOSS (*Free and Open Source Software*) utilizadas en el campo de las humanidades digitales (HD) comprende:

1. *Content Management Systems*:

- *MediaWiki*
- *Drupal*
- *BackdropCMS*
- *WissKI*

- Mukurtu
- Omeka Classic
- Omeka S
- AlchemyCMS

2. Digital Assets Management:

- Islandora
- Arches
- CollectiveAccess

3. Virtual Research Environment:

- ResearchSpace
- Nodegoat

4. Crowdsourcing:

- Zooniverse
- Pybossa
- EnrichEuropeana
- CrowdHeritage

Como veremos a continuación, todas las aplicaciones seleccionadas proporcionan al desarrollador un entorno de trabajo preparado para la configuración, con un mayor o menor margen de customización.

3.1. Content Management Systems (CMS)

Tabla 1. Características básicas de los CMS seleccionados

Categoría	Software	Versión analizada	Actualizada el	Lanzamiento	Lenguajes de programación	Desarrollador	Licencia
CMS	Media Wiki	1.35	25/9/20	2003	PHP, JavaScript	Fundación Wikimedia	GPLv2+
	Drupal	7.77	03/12/20	2011*	PHP, JavaScript	Asociación Drupal	GPLv2/v3
	Drupal	8.9.11	03/12/20	2015*	PHP, JavaScript	Asociación Drupal	GPLv2/v3
	BackdropCMS	1.17.4	26/11/20	2015	PHP, JavaScript	Backdrop PMC	GPLv2+
	Mukurtu	2.1.6	23/11/20	2012	PHP, JavaScript	CDSC - WSU	GPLv3
	WissKI	8.x-2.3	06/11/20	2012	PHP, JavaScript	FAU-GNM-ZFMK	GPLv2+
	Omeka Classic	2.6.1	02/05/18	2008	PHP	CHNM - GMU	GPL
	Omeka S	3.0.1	20/10/20	2019	PHP, JavaScript	CHNM - GMU	GPL
	AlchemyCMS	5.0	17/07/20	2010	Ruby	C. Fregin, T. von Deyen	BDS

* Para Drupal indicamos la fecha de publicación de cada versión en vez que la fecha inicial de lanzamiento del programa, que fue 2001.

MediaWiki es el software de base de Wikipedia y de otros proyectos de la Fundación Wikimedia. Se lanzó en 2003 y desde entonces ha tenido una gran expansión y se ha utilizado en un gran número de wikis, con contenidos muy diferentes. Su uso en proyectos de HD es bastante extendido, por ejemplo para actividades de transcripción colaborativa (Causer; Wallace, 2012), catálogos de bienes patrimoniales (Chevalier et al., 2012; Care, 2020), corpus de textos (Rutherford et al., 2018) y todo tipo de colecciones web, en particular a través de su extensión Semantic MediaWiki, que añade funciones para el manejo de datos estructurados, el soporte para SQL y SPARQL y el control granular de permisos.

Drupal es una plataforma software nacida en 2001: modular, multipropósito y altamente configurable. Es el tercer CMS más popular⁴, con una cuota de mercado del 4,7%. Con respecto a sus competidores, de los que se diferencia por una más pronunciada curva de aprendizaje y una mayor flexibilidad, se considera más bien un CMF (Content Management Framework), es decir un entorno para el desarrollo de aplicaciones web personalizadas. En las últimas versiones⁵ integra por defecto funciones para el manejo y la consulta de datos estructurados. Sus puntos de fuerza son varios:

- el alto número de extensiones disponibles (módulos⁶), que en algunos casos cubren necesidades muy relevantes en HD⁷;
- la posibilidad de combinar e integrar módulos para la creación de funciones adicionales;
- sus sólidas capacidades de mapeo y geocodificación;
- una amplísima comunidad de desarrolladores y usuarios.

Su uso en proyectos HD ha sido muy extendido (Dombrowski, 2016a; 2016b; Velios; Martin, 2017), sobre todo en su versión 7, lanzada en 2011 y cuyo soporte oficial ha sido recientemente extendido hasta finales de 2023. El lanzamiento

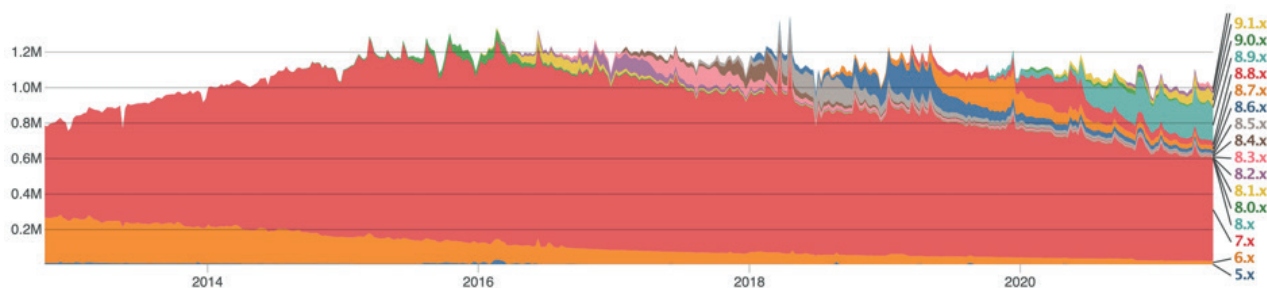


Figura 2. Grado de adopción y uso de las versiones de *Drupal*, desde la 5 hasta la 9.1. Se puede apreciar como el uso de la versión 7, con el 66% del total, sigue siendo prevalente.

Fuente: <https://www.drupal.org/project/usage/drupal>

de su versión 8, en 2015, ha introducido unos cambios relevantes en la arquitectura de base⁸, orientados a cubrir sobre todo las necesidades de grandes empresas, lo que ha incrementado el coste de desarrollo y los requisitos de alojamiento web. En general, esos cambios han frenado el grado de adopción de *Drupal*⁹ (figura 2) y, en especial, su uso en proyectos de HD (Dombrowski, 2018), y han supuesto el nacimiento, en enero de 2015, de un *fork* derivado: *BackdropCMS*. La versión 9, que mantiene una marcada continuidad con la versión 8, fue lanzada el 3 de junio de 2020, pero su uso es aún marginal.

Como hemos señalado, *BackdropCMS* nace como un *fork* de *Drupal*, del cual ha mejorado muchas de las deficiencias bajo los principios de personalización, asequibilidad, rapidez, retro-compatibilidad y extensibilidad. En la hoja de ruta de su desarrollo pone mucho énfasis en la experiencia a nivel editorial (flujo de trabajo) y en la flexibilidad de la arquitectura de la información, lo que, unido con la herencia de *Drupal 7* en cuanto a modelación de datos y creación de vistas dinámicas, lo hace particularmente atractivo en el contexto de las HD y para la implementación de SIW personalizados. Su mayor desventaja es el nivel de adopción¹⁰, aún escaso si se compara con *Drupal*, y la baja disponibilidad de módulos¹¹.

Un fenómeno interesante en el contexto de *Drupal* son las distribuciones¹²: versiones pre-configuradas que unen un conjunto de módulos para satisfacer las necesidades de una aplicación específica, algunos de los cuales muy relevantes para nuestro contexto. Se trata, en cierto modo, de un nuevo software que utiliza *Drupal* como componente de base. Dos de las distribuciones más interesantes para proyectos en patrimonio cultural y HD son *Mukurtu* y *WissKI*.

Mukurtu es un sistema de gestión de contenidos pensado para proporcionar acceso digital al patrimonio cultural, co-diseñado para satisfacer las particulares necesidades de las comunidades indígenas para gestionar, dar a conocer y compartir su patrimonio a nivel digital, de manera culturalmente relevante y ética. Ha sido desarrollado y es mantenido por el *Center for Digital Scholarship and Curation* de la *Washington State University* como una plataforma FOSS. *Mukurtu* está orientado en particular hacia la curación y el uso de metadatos para describir y etiquetar correctamente el patrimonio indígena de una comunidad, proporcionando un acceso diferencial a la información para los miembros de la comunidad y para el público en general, a través de los que denomina “protocolos culturales” (Wilberg, 2014). Su plan de desarrollo futuro, que es público y fue actualizado a finales de 2020¹³, conlleva el pasaje a *Drupal* con la versión 4 de *Mukurtu*, prevista para finales de 2021, y un periodo de migración de un año para que los usuarios puedan migrar a esa nueva versión antes que finalice el soporte oficial a *Drupal*.

WissKI es definido por sus creadores como un entorno virtual de investigación basado en la web (VRE) y un sistema de gestión de contenidos (CMS) que, a pesar de ser multiusuario, mantiene un enfoque explícito en el ámbito de las HD¹⁴. Ha sido producido por tres instituciones científicas:

- el Grupo de Investigación de Humanidades Digitales del Departamento de Ciencias Informáticas de la Universidad Friedrich Alexander de Erlangen-Nuremberg (FAU);
- el Departamento de Informática de Museos del Germanisches Nationalmuseum (GNM), y el
- Grupo de Informática de la Biodiversidad del Zoologisches Forschungsmuseum Alexander Koenig (ZFMK).

Implementa conceptos de web semántica y se orienta claramente hacia proyectos científicos que recogen, almacenan, gestionan y comunican conocimiento relacionado con resultados de investigación, con la mirada puesta en la interoperabilidad a largo plazo (Scholz; Goertz, 2012). Incorpora la ontología *Erlangen CRM*, que es una implementación OWL-DL 1.0 (*Web Ontology Language*) de *Cidoc-CRM*, pero soporta la importación de cualquier otra ontología. Como almacén de RDF utiliza el *triplestore ARC14*, con un punto SPARQL para la consulta de datos (Nishanbaev et al., 2019).

Los datos, en particular en las disciplinas que estudian el pasado, no suelen perder actualidad con el tiempo; al contrario, los sistemas de información web (SIW) pertenecen a un entorno en continua evolución, estrechamente dependiente de las tecnologías disponibles y empleadas en un determinado momento

Desde la perspectiva de las HD, *Omeka* se considera una de las principales alternativas a *Drupal* y sus distribuciones, pero como veremos a continuación su estructura está diseñada sobre todo para gestionar colecciones digitales de objetos enriquecidos con metadatos. Ha sido llevada a cabo a partir de 2007 por el *Roy Rosenzweig Center for History and New Media* de la *George Mason University* (Virginia), como una solución software para la catalogación y presentación de objetos culturales procedentes de bibliotecas, museos, archivos, colecciones y exposiciones virtuales (Alcaraz-Martínez, 2012). Se encuentra disponible en dos versiones: *Omeka Classic* y *Omeka S*. La primera, orientada hacia proyectos individuales y el sector educativo, está basada en el estándar de metadatos *Dublin Core* y sus funciones pueden ampliarse a través de *plugins*. Las posibilidades de extensión son mucho más limitadas que en *Drupal*, pero existen módulos muy relevantes para nuestro contexto de interés, tales como *Neatline*, para incorporar mapas y líneas temporales, y *Scripto*, una herramienta para las transcripciones colaborativas. *Omeka S*, cuya versión 1.0 fue publicada a finales de 2017, está orientada hacia un ámbito más institucional, como un recurso centralizado desde el que desarrollar múltiples plataformas. Además, no está limitada a *Dublin Core*, sino que puede implementar múltiples vocabularios/ontologías y facilita la vinculación de datos abiertos enlazados. En ambas versiones, la arquitectura del modelo de datos basada en estándares de metadatos constituye la principal diferencia con *Drupal*, donde el modelo de datos es completamente personalizable. Además de almacenar los items directamente en la base de datos de *Omeka*, existe la posibilidad de importar objetos digitales (archivos XML, como TEI o EAD, imágenes o videos) desde un repositorio *Fedora Commons* como datos enlazados, a través del módulo *Fedora Connector* (Gruber, 2020).

En esta panorámica merece una mención el software *AlchemyCMS*. Se trata del sistema de gestión de contenidos utilizado por *Europeana Virtual Exhibitions*. A nivel informático es una plataforma *headless*, es decir que proporciona solo un *back-end* construido como un repositorio de contenido estructurado, que suministra a través de una RESTful API. Por su propia arquitectura, se centra en almacenar contenido puro, sin lenguaje de marcado, estilo o formato. No proporciona una plataforma lista para usarse, pero pone a disposición los componentes necesarios para desarrollar un CMS a medida. Integra la tecnología *ImageMagick* y por eso es particularmente apta para la creación de sistemas basados en imágenes. El proyecto empezó como un aplicativo propietario en 2007, bajo el nombre *washAPP*, y se convirtió a código abierto en 2010. Ofrece una documentación muy amplia y estructurada y cuenta con una sólida y activa base de desarrolladores.

3.2. Digital Assets Management (DAM)

Tabla 2. Características básicas de los DAM seleccionados

Categoría	Software	Versión analizada	Actualizada el	Lanzamiento	Lenguajes de programación	Desarrollador	Licencia
DAM	<i>Islandora</i>	7.x-1.13	03/09/19	2006	PHP, Java, JavaScript	<i>Islandora Foundation</i>	GPL
	<i>Arches</i>	5.1	27/10/20	2013	Python, Javascript	<i>Getty CI - World Monuments Fund</i>	GPLv3
	<i>Collective Access</i>	1.7.9	01/01/21	2009	PHP, JavaScript	<i>Whirl-i-Gig</i>	GPLv3

Pasando a la categoría de sistemas de gestión de activos digitales, *Islandora* representa una solución de software bien consolidada para la creación de colecciones digitales y la conservación a largo plazo de los recursos. Procede desde la combinación de *Fedora* y *Drupal*, para ayudar instituciones y organizaciones a gestionar y localizar colecciones de objetos digitales de manera colaborativa con su público. Originalmente desarrollado por la *University of Prince Edward Island's Robertson Library*, es ahora mantenido por una fundación internacional. Utiliza *Fedora* como repositorio para el almacenamiento de los objetos digitales, el control de acceso y el *versioning*; *Drupal* para la creación de la interfaz web, la interacción con los usuarios, la creación de los metadatos, las taxonomías y la expansibilidad a través de módulos¹⁵. *Islandora* añade a la combinación entre *Fedora* y *Drupal* toda una serie de módulos específicos para la configuración de la plataforma y la interoperabilidad con servicios externos (microservicios y servicios de almacenamiento). Puede almacenar y visualizar una muy amplia variedad de objetos: imágenes, audios, videos, PDFs, colecciones, contenido paginado (libros, periódicos, seriales) y archivos binarios (Ruest; Stapelfeldt, 2014; Stapelfeldt; Moses, 2013).

El software *Collective Access* empezó a gestarse en 2003 bajo el nombre *OpenCollection* por la compañía *Whirl-i-Gig*, que lanzó su primera versión pública en 2007 (Alcaraz-Martínez, 2014). Está pensado para administrar y publicar en la Web grandes y heterogéneas colecciones de bibliotecas, museos y archivos, que requieren soporte para diferentes estándares de metadatos (*Dublin Core*, *PBCore* y *VRA Core* vienen integrados), tesauros de referencia (*Library of Congress subject headings*, *Getty*, *GeoNames*), formatos de medios (imágenes, video, audio, documentos multi-páginas, escaneos 3D) y una variedad de opciones de visualización (zoom teselado para imágenes en alta resolución, anotaciones, líneas dinámicas del tiempo, mapas, etc.). A nivel de infraestructura informática, *Collective Access* se basa en dos componentes principales: *Providence*, como interfaz de catalogación *back end*, y *Pawtucket*, como plataforma de consulta y publicación *front end*.

Volviendo al campo más específico de la gestión del patrimonio cultural, una solución muy completa es *Arches*, desarrollada conjuntamente por el *Getty Conservation Institute (GCI)* y el *World Monuments Fund (WMF)*, con una amplia participación de profesionales del patrimonio a nivel internacional. Está pensada para un uso prevalentemente institu-

cional, con el fin de crear inventarios digitales que describan tipo, ubicación, extensión, períodos culturales, materiales y condiciones de los recursos patrimoniales y permitan establecer relaciones entre ellos. La plataforma *Arches* incluye: un sistema para gestionar, definir y estructurar los datos; varios sistemas de consulta y visualización para buscar, señalar y representar elementos (también a nivel espacial, temporal y *fuzzy*); un gestor

de proyectos y tareas, para organizar la edición de datos, estableciendo flujos de trabajo personalizados. Soporta el uso de múltiples ontologías, a partir de *Cidoc-CRM*, y la definición de vocabularios controlados (Carlisle et al., 2014). Además, sigue los estándares del *Open Geospatial Consortium (OGC)*, lo que significa que la plataforma es compatible con las aplicaciones GIS de escritorio. Desde la versión 4, ofrece la aplicación móvil *Arches Collector*, destinada a proyectos de recopilación y edición de datos en campo. Recientemente, las capacidades básicas y las tecnologías subyacentes de *Arches* se han ampliado para abarcar una más amplia gama de casos de uso¹⁶, por ejemplo, a través de la iniciativa “*Arches for science*”, cuya versión 1.0 está prevista para el verano de 2021, con el propósito de soportar la comunidad de investigadores en el campo de la ciencia y conservación del patrimonio (*heritage science*).

En relación al público objetivo podemos identificar dos tendencias: software desarrollado para clientes institucionales versus paquetes más adecuados para proyectos individuales, con estructura y funciones personalizadas

3.3. Virtual Research Environments (VRE)

Tabla 3. Características básicas de los VRE seleccionados

Categoría	Software	Versión analizada	Actualizada el	Lanzamiento	Lenguajes de programación	Desarrollador	Licencia
VRE	<i>Research Space</i>	3.4.0	20/02/20	2018	JavaScript, TypeScript, Java	<i>The British Museum</i>	GPLv3+
	<i>Nodegoat</i>	7.3	17/09/19	2013	PHP, JavaScript, C++	<i>Lab1100</i>	GPLv3

ResearchSpace es un entorno virtual de investigación (VRE) desarrollado por el *British Museum* en colaboración con la empresa alemana *Metaphacts* y se encuentra actualmente en fase de *pre-release*. Está focalizado en modelos de datos basado en grafos y emplea tecnologías de la web semántica para apoyar la investigación en patrimonio cultural. Pone el énfasis en la representación gráfica del conocimiento, más que en la catalogación, y promueve la investigación colaborativa (Oldman; Tanase, 2018). Proporciona un entorno integrado de programas diseñados para reflejar los métodos de investigación, incluyendo búsqueda semántica basada en *Cidoc-CRM*, portapapeles y anotaciones semánticas, herramientas de manipulación y anotación de imágenes utilizando *IIF (International Image Interoperability Framework)* y mapas del conocimiento. A pesar de encontrarse aún en un estado temprano de desarrollo, se ha documentado su uso en varias disciplinas, por ejemplo: en arqueología, el proyecto *Gravitate*, focalizado en ofrecer una solución informática para establecer relaciones, culturales o físicas, entre objetos patrimoniales; en historia del arte, el proyecto *Late Hokusai*; en historia de la ciencia, el proyecto *CorpusTracer* (Kräutli; Valleriani, 2018).

Nodegoat comparte con *ResearchSpace* el enfoque orientado a la visualización dinámica de la información y a la generación de conocimiento. Se trata de un entorno virtual de investigación implementado desde 2011 por *LAB1100*, una spin-off de la *Universidad de Amsterdam*. Incluye herramientas de gestión de datos, de análisis de redes y de visualización gráfica. *Nodegoat* está directamente pensado para la investigación, más que para la clasificación y preservación de contenido. Permite a los usuarios co-crear conjuntos de datos basados en modelos diseñados a medida. En el mismo entorno de gestión, ofrece módulos para analizar y visualizar esos conjuntos de datos de forma relacional, diacrónica y espacial (Bree; Kessels, 2013). Ha sido utilizado en varias disciplinas humanísticas para una amplia pluralidad de casos¹⁷, en particular en proyectos donde las preguntas de investigación estaban muy ligadas a la combinación de las variables de espacio y tiempo. Una versión gratuita de *Nodegoat*, pero limitada, está disponible en modalidad *software como un servicio (SaaS)*, mientras que las otras opciones de uso prevén la suscripción de un contrato de servicio anual con *LAB1100*. Existe la posibilidad de descarga libre de todo el código fuente, pero más bien con el fin de garantizar la transparencia del proyecto que para un uso independiente.

3.4. Crowdsourcing

El último grupo de software que pasamos a analizar reúne soluciones informáticas orientadas hacia la creación de plataformas para *crowdsourcing* y la ciencia ciudadana. Hemos seleccionado cuatro: *Zooniverse*, *Pybossa*, *EnrichEuropeana* y *CrowdHeritage*.

Zooniverse es un portal nacido en 2009 desde el proyecto colaborativo *Galaxy Zoo*. No se limita a las humanidades, sino que abraza proyectos de cualquier tipo de la investigación científica. Está basado en la colaboración abierta entre pares y ofrece la posibilidad de crear proyectos para la recolección y manipulación de datos involucrando la ciudadanía. Cuenta con una muy amplia comunidad de usuarios, que en 2019 llegaba a 1,6 millones. A pesar de publicar buena parte de su código en abierto, está pensado para usarse como un servicio a través de su plataforma web. Ofrece herramientas *ad*

Tabla 4. Características básicas de las plataformas para *crowdsourcing*

Categoría	Software	Versión analizada	Actualizada el	Lanzamiento	Lenguajes de programación	Desarrollador	Licencia
Crowdsourcing	Zooniverse	online	20/12/20	2009	JavaScript, Ruby	Fingerprint Digital Media	Apache 2.0
	Pybossa	3.3.0	20/12/20	2013	Python	Scifabric	GPLv3+
	EnrichEuropeana	1.0	01/12/20	2019	PHP, JavaScript, Java	EnrichEuropeana project	GPLv2/EUPLV.1.2
	CrowdHeritage	1.0	05/12/20	2019	JavaScript	ALLS-Lab	EUPL V.1.2

hoc para asignar tareas a los participantes, así como la posibilidad de cargar conjuntos de datos (normalmente imágenes) sobre los que se pide a los voluntarios que realicen operaciones; la estructura de su modelo de datos no se puede adaptar a casos específicos (Simpson et al., 2014).

Pybossa es un *framework* de código abierto para implementar plataformas de recogida, análisis y enriquecimiento de datos en entornos colaborativos, lanzado en 2013 por *Scifabric*, una compañía con sede en España que sigue ocupándose de su desarrollo. Ofrece funciones de transcripción de documentos manuscritos; transcripción de vídeo y audio; análisis de objetos multimedia; geolocalización de archivos; identificación y etiquetado de objetos, caras, sonidos, etc., en imágenes, audio y vídeo; identificación de secciones de obras digitalizadas, tales como animales, personas, análisis de sentimientos, etc.; enriquecimiento de registros y archivos con metadatos. Su aplicación no se limita a proyectos de ciencia ciudadana, pero muchos de sus casos de uso proceden de las humanidades. En particular, hay que señalar que la tecnología ofrecida por *Pybossa* constituye la infraestructura base de la iniciativa *Comunidad BNE*, promovida en 2019 por la *Biblioteca Nacional de España* en colaboración con *Red.es* (Sánchez-Nogales, 2019). *Comunidad BNE* es una plataforma de trabajo colaborativo, abierto y distribuido, sobre las colecciones digitales y datos de la *Biblioteca Nacional de España*, similar a *Zooniverse*, pero con dos diferencias importantes: los investigadores no pueden aportar sus propios datos; se pueden sugerir nuevos proyectos, pero no crearlos directamente. A los usuarios se le pueden asignar tareas de identificación de objetos, transcripción de textos y audios, georreferenciación, marcado y etiquetado, enriquecimiento de metadatos y corrección OCR. *BNElab*, que está a cargo de su desarrollo, ha añadido nuevas funciones a la tecnología de base que, a su vez, ha puesto a disposición para su libre reutilización.

EnrichEuropeana es el nombre de un paquete software desarrollado en el marco del homónimo proyecto, financiado por *Europeana* y la *Comisión Europea* entre 2018 y 2019 con el fin de renovar la infraestructura tecnológica de la plataforma *Transcribathon* (Sciotti, 2019). Como otras plataformas para el *crowdsourcing*, ofrece funciones para transcribir, anotar y georreferenciar documentos históricos, pero tiene un ámbito de aplicación más restringido que las otras descritas anteriormente. A nivel tecnológico, está compuesta por tres módulos: una aplicación *front end* compuesta por una versión personalizada de *WordPress* y un visor *IIIF*; una API para el enriquecimiento semántico; una infraestructura de intercambio de datos para la transferencia de información entre *Europeana* y la plataforma *Transcribathon*.

La plataforma *CrowdHeritage* es la última iniciativa en orden de tiempo promovida por *Europeana* en este campo. Fue creada con el fin de utilizar *crowdsourcing* para mejorar la calidad de los metadatos del patrimonio cultural digital indexado en *Europeana*, así como en otras bases de datos de instituciones dedicadas al patrimonio que cuenten con APIs públicas. Ha sido desarrollada por la *Universidad Nacional Técnica de Atenas*, en colaboración con la *European Fashion Heritage Association*, la *Michael Culture Association*, el ministerio francés de la cultura y la *Fundación Europea* (Kaldeli et al., 2021). A diferencia de otras soluciones informáticas en este campo, el tipo de aplicaciones que pone a disposición es algo más restringido, focalizándose en anotación del contenido, marcado semántico a través de tesauros y vocabularios, análisis de colores y geo-tagging colaborativo. La participación de los usuarios se promueve a través de la gamificación, utilizando una tabla de clasificación y asignando insignias. En principio, todo el código está disponible en abierto y puede ser utilizado para implementar plataformas similares, aunque no sea ese el objetivo de la iniciativa.

Los *virtual research environments* (VRE) se caracterizan por un elevado grado de personalización en el diseño del modelo de datos y en la definición de las taxonomías, junto a una escasa o nula posibilidad de adaptación de las funciones

4. Análisis comparativo

Una vez realizada la descripción individual de las aplicaciones seleccionadas, en las tablas 5-7 mostramos una comparativa de sus principales características, divididas en tres categorías:

- funciones esenciales;
- estructuración y recuperación de la información;
- funciones adicionales.

En el conjunto, hemos elegido hasta 15 propiedades, entre más básicas y más avanzadas, que suelen ser necesarias para la creación de sistemas de información web y que los creadores/usuarios de estos sistemas suelen requerir.

El riesgo intrínseco en la elaboración de tablas comparativas es exceder en la simplificación a beneficio de la visión de conjunto. En muchos casos, la calidad y la exhaustividad en la implementación de una cierta componente son muy distintas en un software respecto a otro y señalar la simple presencia/ausencia puede resultar reductivo. Para compensar, en parte, esta limitación hemos introducido dos variantes: “L” por limitado, para indicar los casos de una implementación parcial o experimental; “E” por extensión, si la funcionalidad depende de un paquete externo (módulo o *plugin*), lo que de por sí no habla de su calidad, pero sí de su nivel de integración y perspectiva de mantenimiento.

Tabla 5. Comparativa de las funciones esenciales de las plataformas analizadas

Software	Formatos básicos*	Formatos avanzados**	Mapas	Multi-lenguaje	Importación / exportación	Extensible con módulos
MediaWiki	X	E	E	X	X	X
Drupal 7	E	E	E	X	E	X
Drupal 8/9	X	E	E	X	E	X
BackdropCMS	X		L	X	X	X
WissKI	X	E	E	X***	X	X
Mukurtu	X	E	X	X***	X	X
Omeka Classic	X	E	L	X	X	X
Omeka S	X	E	L	X	X	X
AlchemyCMS	X			L		L
Islandora	X	X	X	X***	X	X
Arches	X	X	X	X	X	X
CollectiveAccess	X	X	X	X	X	
ResearchSpace	X		X		X	
Nodegoat	X		X		X	
Zooniverse	X		X	L	X	
Pybossa	X	X	X	L	X	X
EnrichEuropeana	X					
CrowdHeritage	X			X		

Leyenda: E = extensión; L = limitado.

* PDFs, imágenes, audios y videos. ** Objetos 3D, colecciones, contenido paginado, binarios, KML, etc.

*** Soporte multilinguaje basado en *Drupal*, pero no disponible para la interfaz de la propia distribución.

Hemos dividido la gestión de diferentes formatos entre básicos y avanzados, ya que prácticamente todos incluyen los primeros por defecto, mientras que la gestión de formatos más avanzados es garantizada solo en el caso de los DAMs, ya en el *core*, y en la mayoría de los CMS, a través de extensiones. Las funciones de mapeo, que poseen una indudable importancia para el estudio del pasado, tradicionalmente en disciplinas como la Arqueología y el estudio del Patrimonio, y más recientemente en Historia del Arte y Filología, están cubiertas en la gran mayoría de los casos, aunque a veces de manera limitada. Situación similar a la observada para la función multilinguaje, que hemos incluido entre las esenciales por la relevancia que el multilingüismo y la diversidad lingüística han adquirido en el campo de las HD (Budín, 2015)¹⁸.

Las carencias más evidentes se encuentran entre los VRE y las plataformas para el *crowdsourcing*, con la excepción de *Pybossa*, y se pueden explicar por ser soluciones para un público más específico y un enfoque más acotado, que se limitan a las funciones consideradas necesarias en su campo de aplicación y no prevén mucha extensibilidad.

Existe claramente una tensión en el modelado de datos entre estandarización, compatibilidad y interoperabilidad, por un lado, y personalización y peculiaridades, por otro, con argumentos válidos a favor de ambas estrategias (González-Pérez, 2018). Una representación de estas tendencias opuestas se muestra en las dos primeras columnas de la tabla 6, donde se puede identificar algún patrón. *Drupal* y *BackdropCMS* favorecen claramente la

La recomendación de utilizar paquetes software existentes se basa en la preferencia de utilizar código común para implementar sistemas de información web (SIW) personalizados en lugar de código personalizado para SIW que comparten características y funciones comunes

Tabla 6. Comparativa de las funciones de estructuración y recuperación de la información

Software	Modelo de datos personalizable	Estándares de metadatos	Datos enlazados	Búsqueda básica	Búsqueda avanzada
MediaWiki	E	E	E	X	Elasticsearch
Drupal 7	X	L		X	Solr
Drupal 8/9	X	L		X	Solr
BackdropCMS	X			X	
WissKI	X	X	X	X	Solr
Mukurtu	X	X	X	X	Solr
Omeka Classic	L	X		X	
Omeka S	L	X	X	X	Solr
AlchemyCMS	L			X	Ferret
Islandora	X	X	X	Solr	Solr
Arches	X	X	X	X	Elasticsearch
Collective Access	X	X		X	X
ResearchSpace	X	X	X	X	X
Nodegoat	X	X	X	X	X
Zooniverse				X	
Pybossa	L			X	
EnrichEuropeana		X	L	X	
CrowdHeritage		X	X	X	

Leyenda: E = extensión; L = limitado.

personalización frente al soporte de los estándares de metadatos, mientras que las aplicaciones basadas en *Omeka* y *Europeana* hacen lo contrario; varias soluciones ofrecen un buen equilibrio de ambas características, y algunas de ellas añaden la interoperabilidad con los datos enlazados (Davis; Heravi, 2021).

La búsqueda básica está incluida en todas las plataformas, y la mayoría de ellas ofrecen también soporte para funciones avanzadas de recuperación de la información, mediante la integración de un motor de búsqueda externo especializado. *Apache Solr* destaca claramente como la solución preferida, pero existen alternativas. Todas ellas ofrecen un mejor rendimiento y una buena escalabilidad para conjuntos de datos muy grandes. La falta de capacidades de búsqueda avanzada en el grupo de software dedicado al *crowdsourcing* puede explicarse por el hecho de que, por diseño, prefieren que los usuarios naveguen por los contenidos siguiendo un flujo de trabajo preestablecido, en lugar de buscar libremente, con el fin de optimizar la realización de las tareas asignadas.

El último conjunto de características seleccionadas para la comparación no es esencial para la implementación de un SIW, pero ofrece valiosas opciones adicionales. La opción SaaS está disponible, a veces como opción por defecto, para todo el software dedicado al *crowdsourcing*. Esto es razonable, ya que esas soluciones responden a flujos de trabajo específicos que pueden gestionarse de forma centralizada y desplegarse en varias instancias, que difieren en el contenido, pero no en las funciones o la estructura de datos.

No hay que subestimar la disponibilidad de una aplicación móvil oficial, tanto para mejorar la experiencia de navegación de grandes conjuntos de datos en dispositivos móviles (tabletas en particular) como para proyectos que impliquen tareas de recogida de datos que requieran acceso a componentes hardware (cámara, GPS, giroscopio, etc.). Muy pocas plataformas disponen de una, pero la mayoría del resto proporciona una API REST, por lo que se pueden desarrollar aplicaciones móviles personalizadas con relativa facilidad.

La disponibilidad de versiones demo online no es exactamente una funcionalidad del software y son menos relevantes cuando se dispone de una buena documentación o de un servicio SaaS, pero pueden ayudar a evaluar los productos más complejos de instalar. El soporte oficial, con una combinación de servicios de pago y gratuitos, es básicamente disponible para todas las plataformas seleccionadas.

A nivel de conjunto, una primera distinción necesaria hay que realizarla entre 1) plataformas agnósticas y 2) desarrolladas para una aplicación específica.

1) Pertenecen claramente al primer grupo:

MediaWiki, Drupal, Islandora, CollectiveAccess, AlchemyCMS, Zooniverse y Pybossa.

2) Están pensadas para un campo de aplicación más concreto:

BackdropCMS (pequeñas empresas y ONGs); *WissKI*, *Nodegoat* y *Research Space* (investigación en HD); *Mukurtu* (patrimonio cultural de comunidades indígenas); *Omeka* (colecciones digitales); *Arches* (inventarios patrimoniales); *EnrichEuropeana* y *CrowdHeritage* (transcripción y etiquetado de colecciones indexadas en *Europeana*).

Otra variable que afecta significativamente las características de algunas plataformas frente a otras es el público objetivo. Podemos identificar dos tendencias principales

1) Un primer grupo de software (al que pertenecen *CollectiveAccess*, *Omeka S* y *Arches*) ha sido desarrollado pensando sobre todo para clientes institucionales, como los GLAM: se trata de soluciones orientadas a la creación de repositorios institucionales que abogan por una clara preferencia hacia los modelos de datos basados en metadatos estandarizados, la interoperabilidad (datos enlazados y API) y la adopción de protocolos comunes.

2) Un segundo grupo es más adecuado para implementar soluciones web para proyectos individuales, con estructura y funciones personalizadas. Algunos paquetes, como *WissKI*, *Islandora* o *Nodegoat*, ofrecen un enfoque híbrido, que combina un buen nivel de compatibilidad e interoperabilidad con estándares consolidados, como *Dublin Core* o *Cidoc-CRM*, con la posibilidad de personalizar el modelo de datos y desarrollar ontologías personalizadas a partir de estándares comunes. Como consecuencia de esta ampliación de funciones, se ha incrementado también el peso y la complejidad del código fuente, así como el número de dependencias externas, y el panorama ofrece ya pocas opciones de fácil instalación y con unos requisitos básicos en cuanto a alojamiento web. Entre estas últimas podemos nombrar: *MediaWiki*, *Drupal*, *BackdropCMS*, *Omeka Classic* y *Mukurtu*.

En general, son todas soluciones muy maduras, con trayectorias de desarrollo relativamente largas (tablas 1-4). Solo dos, *ResearchSpace* y *CrowdHeritage*, son aplicativos nuevos aparecidos en los últimos cinco años, ya que *Omeka S*, *BackdropCMS* y *EnrichEuropeana* son evoluciones de plataformas preexistentes. El nivel de adopción es muy variado y difícilmente comparable, sobre todo entre aplicaciones agnósticas y especializadas.

Finalmente, es necesario realizar una última distinción entre los proyectos que publican su código fuente en abierto para favorecer su distribución, descarga y uso, y aquellos que utilizan el *open source* principalmente para asegurar la transparencia de la iniciativa y la estabilidad del código a largo-plazo. En el primer caso, el proceso de instalación suele ser relativamente simple o muy bien documentado y, si la aplicación necesita de dependencias software externas, suelen ser relativamente comunes y ampliamente utilizadas en entornos de desarrollo web. Entran en este grupo la mayoría de las soluciones analizadas en esta reseña, como: *MediaWiki*, *Drupal*, *BackdropCMS*, *WissKI*, *Mukurtu*, *Omeka*, *AlchemyCMS*, *Islandora*, *Arches*, *CollectiveAccess* y *ResearchSpace*. En cambio, al segundo grupo pertenecen aplicaciones que suelen ofrecerse en la modalidad SaaS y que proporcionan apoyo activo a su comunidad de usuarios (ej. *Zooniverse*, *EnrichEuropeana* y *CrowdHeritage*) o al desarrollo de los proyectos que albergan (ej. *Nodegoat* y *Pybossa*).

5. Limitaciones

Una revisión de este tipo sería incompleta sin un apartado dedicado a las limitaciones que afectan esta clase de productos software, siendo la necesidad de mantenimiento sistemático una de las más relevantes. De hecho, ninguna de las soluciones software analizadas en este artículo puede mantenerse con seguridad online sin un mantenimiento correctivo y adaptativo (Sommerville, 2005), que implica tanto recursos técnicos como humanos.

Las instalaciones software¹⁹, especialmente las accesibles en red, se ven afectadas por la identificación y divulgación de fallos de seguridad que pueden exponer a ataques automatizados a gran escala, que comprometen incluso sistemas que no son objetivo directo de ataque. Destacan, en este sentido, los casos conocidos como *Drupalgeddon 1* y *2*, ocu-

Tabla 7. Comparativa de las funciones adicionales

Software	SaaS	App móvil	REST API	Demo oficial
<i>MediaWiki</i>			X	
<i>Drupal 7</i>			E	
<i>Drupal 8/9</i>			X	
<i>BackdropCMS</i>			E	X
<i>WissKI</i>			X	
<i>Mukurtu</i>		X		X
<i>Omeka Classic</i>		X	X	
<i>Omeka S</i>	X	X	X	
<i>AlchemyCMS</i>			X	X
<i>Islandora</i>			X	X
<i>Arches</i>		X		X
<i>CollectiveAccess</i>	X		X	X
<i>ResearchSpace</i>			L	L*
<i>Nodegoat</i>	X		X	X
<i>Zooniverse</i>	X	X		X
<i>Pybossa</i>	X		X	
<i>EnrichEuropeana</i>	L**		X	L**
<i>CrowdHeritage</i>	L**		X	L**

Leyenda: E = extensión; L = limitado.

* Está disponible una versión demo oficial, pero debe ser configurada manualmente con *docker-compose*. ** Disponible como SaaS sólo para sus respectivos proyectos patrocinadores; esas instancias también pueden considerarse software de demostración.

rridos respectivamente en 2014 y en 2018²⁰. Expusieron vulnerabilidades que posibilitaron el hackeo automático, en pocas horas, de prácticamente todas las plataformas web basadas en *Drupal* no actualizadas a la última versión y la creación de *backdoors* que no fueron eliminados por la posterior actualización del software. La mejor protección contra la divulgación pública, y la posterior posible explotación maliciosa, de fallos de seguridad es la pronta instalación de los parches de seguridad desarrollados, en particular los relativos a los componentes *core*, pero que a su vez pueden introducir incompatibilidades con otros componentes del sistema, por lo que es necesario un servicio de mantenimiento activo.

El abanico de soluciones disponibles es lo suficientemente amplio y maduro como para preferir su uso frente a la opción de programar el sistema de información web (SIW) desde cero

Además hay que tener en cuenta que la Web es un espacio tecnológico en continua transformación y actualización, ya sea lado servidor o lado cliente. Esta continua evolución provoca, cada cierto tiempo, incompatibilidades de los sistemas desarrollados con las tecnologías web que sustentan su alojamiento en un servidor. Un ejemplo representativo en este sentido ha sido la evolución del lenguaje PHP, en absoluto el más popular en la web²¹. La transición desde la versión 5, soportada hasta 2018, a la versión 7, publicada en 2015, supuso relevantes mejoras, pero que causaron incompatibilidades con las aplicaciones existentes que obligaron a una actualización manual del código fuente. En la elección de una solución software para desarrollar un SIW, es fundamental tener en cuenta la presencia de una amplia y activa comunidad de usuarios y desarrolladoras, ya que la existencia de la primera crea las condiciones necesarias para que los segundos actúen en la resolución de las incompatibilidades sobrevenidas.

Las infraestructuras académicas centralizadas, respaldadas por inversiones institucionales, son tan adecuadas para prestar estos servicios de mantenimiento como escasas. Especialmente los proyectos individuales, a menudo promovidos por un investigador individual para abordar un tema específico de estudio, llegan inevitablemente con el tiempo a un punto de finalización, mientras que es conveniente mantener online los recursos producidos. Queda fuera del alcance de este artículo revisar las soluciones a largo plazo disponibles para abordar esta cuestión, como la conversión de plataformas dinámicas a sitios web estáticos²² o el uso de CMS *flat-file*²³, pero el humanista digital que promueva el desarrollo de un SIW debe ser consciente de estas cuestiones desde el principio y diseñar un plan con antelación.

6. Conclusiones

Después de valorar las características de las soluciones informáticas seleccionadas, podemos concluir que todas las plataformas analizadas tienen sus ventajas y sus desventajas y es necesario adquirir un conocimiento actualizado del panorama disponible en cada momento con el fin de elegir la opción más adecuada para cubrir las necesidades identificadas en cada caso. En particular, existen soluciones que por sus características polifacéticas pueden cubrir un espectro más amplio de necesidades y permitir el despliegue de sistemas de información web que pueden habilitar nuevos procesos de investigación, incluso no previstos en fase de planteamiento inicial (por ejemplo, en lo referido a los procesos de rescate y verificación de fuentes, enriquecimiento de corpus, mapeo colaborativo, etc.).

Al estado actual, *Drupal*, de por sí o a través de un paquete software que lo incorpora (ej. *WissKI*, *Mukurtu* o *Islandora*), ocupa un nicho muy específico y sigue representando una de las mejores opciones disponibles para nuestro ámbito de aplicación. Su nivel de flexibilidad en cuanto a modelo de datos e incorporación de funciones avanzadas a través de módulos lo hace comparable con soluciones completamente personalizadas, pero a la vez proporciona un *front end* amigable y cuenta con una muy amplia comunidad de soporte. Se adapta tanto a la gestión de conjuntos de datos estructurados de investigación como a la creación de repositorios de objetos digitales, sobre todo en ambientes colaborativos, y en las últimas versiones (desde la 8.5) ha mejorado también la gestión de contenido multimedia, tradicionalmente deficiente. *WissKI* en particular, gracias a su *back end* semántico y su larga trayectoria de desarrollo (2012-2021), es una opción muy válida para proyectos de HD de cierta envergadura. Para proyectos más pequeños y con un presupuesto reducido o ausente de cara al mantenimiento, la alternativa de *BackdropCMS* es muy adecuada y solo queda la duda sobre su sostenibilidad a largo plazo, considerando el relativamente bajo nivel de adopción hasta la fecha.

Fuera de la galaxia *Drupal*, los productos más interesantes en el contexto de las HD son tres:

- *Omeka*, que mantiene en ambas versiones un enfoque preferente para la creación de colecciones digitales con metadatos estándar.
- *Pybossa*, para proyectos de ciencia ciudadana que prevean la involucración de la sociedad a través de la asignación de tareas pre-configuradas.
- *Nodegoat*, por el amplio espectro de funciones pre-configuradas que cubren desde la modelación de datos a la publicación online, pasando por la co-creación, la visualización espaciotemporal y el análisis de redes. Pertenece a una categoría, los entornos virtuales de investigación (*Virtual Research Environment*, VRE), que se está afirmando progresivamente como una solución informática válida para la generación, integración, difusión y conservación del conocimiento en ambientes colaborativos (Blake et al., 2010). Se caracterizan por un elevado grado de personalización en el diseño del modelo de datos y en la definición de los vocabularios/taxonomías, junto a una escasa o nula posibilidad de adaptación de las funciones (ej. flujo de datos, opciones de búsquedas, vistas, etc.), que se consideran más generalizadas y aplicables *tout court* a diferentes áreas disciplinares y dominios de investigación.

Una perspectiva interesante para el sector, en particular para proyectos de mayor envergadura, resultaría del uso combinado de repositorios y entornos virtuales de investigación: siendo los primeros el lugar adecuado para el almacenamiento a largo plazo y el intercambio de conjuntos de datos versionados y los segundos el espacio para el procesamiento dinámico y el enriquecimiento con metadatos, en un ambiente eminentemente colaborativo.

Para un público más institucional, no podemos evitar de poner en evidencia la plataforma *Arches*²⁴, que con las últimas versiones ha ampliado su aplicación más allá de la gestión del patrimonio cultural. Presenta una aplicación móvil dedicada al trabajo de campo, unas funciones muy sólidas en cuanto a georreferenciación y mapeo, control granular de los permisos de acceso, gestión integral y flexible del flujo de trabajo, y mantiene pública una hoja de ruta bien definida en cuanto a evolución en los próximos años.

Una conclusión adicional que podemos derivar de este análisis es que el abanico de soluciones disponibles es lo suficientemente amplio y maduro como para preferir, en general, su uso frente a la opción de programar el SIW desde cero, combinando un Sistema de Gestión de Bases de Datos (SGBD) con *backend* y *frontend* personalizados (por ejemplo, el modelo *LAMP*²⁵). Por un lado, este enfoque permite el grado más alto de personalización y proporciona un mayor nivel de seguridad frente a ataques externos, que representan, como se ha visto anteriormente, un factor importante a tener en cuenta utilizando paquetes preconstruidos. Por otro lado, requiere de competencias más especializadas para la implementación del sistema, lo que se traduce en un coste más alto, en particular para el desarrollo inicial, pero también para el simple mantenimiento adaptativo a medio-largo plazo. Expone, además, a un mayor riesgo frente al denominado *bus factor*, ya que la completa personalización de un sistema conlleva un grado más alto de dependencia con respecto al desarrollador inicial y condiciona la posibilidad de transferencia del mantenimiento/ampliación del mismo a otra persona o grupo. Otro coste añadido resulta de la necesidad de producir y mantener documentación sobre el código generado y de hacerse cargo de asegurar la compatibilidad del sistema con librerías externas y tecnologías lado servidor, que inevitablemente se irían actualizando en el tiempo. En definitiva, hemos elegido, en el contexto de esta revisión, limitarnos al uso de soluciones software existentes, pero altamente adaptables. El principio que subyace a nuestra recomendación de utilizar soluciones existentes se basa en la preferencia de utilizar código común para implementar SIW personalizados en lugar de código personalizado para SIW que comparten, en muchos casos, características y funciones comunes.

7. Notas

1. Este criterio nos ha llevado a excluir, por ejemplo, el *Archaeological Recording Kit (ARK)*: un paquete software de código abierto para la gestión de datos en Arqueología. Es la tecnología de base utilizada en la plataforma *Fasti Online* <http://www.fastionline.org>

una base de datos online de excavaciones arqueológicas, prospecciones y proyectos de conservación arqueológica. A pesar de esto, la última versión estable (v.1.1):

<https://sourceforge.net/projects/arkdb/files>

ha sido publicada en 2014 y no es compatible con las actuales versiones de PHP. La versión 2.0 se encuentra en desarrollo desde hace tiempo, pero el repositorio no presenta cambios en los últimos cuatro años:

<https://github.com/lparchaeology/ark2>

2. Se han consultado varias fuentes para identificar proyectos e iniciativas que emplean el software seleccionado para desarrollar SIW y repositorios digitales. Ejemplos de grandes colecciones temáticas disponibles online pueden consultarse en:

https://www.semantic-mediawiki.org/wiki/Projects_in_eHumanities_running_Semantic_MediaWiki

<https://digibug.ugr.es/handle/10481/55656>

<https://doi.org/10.5281/zenodo.3893546>

<https://cohistoria.es/proyectos>

<https://red.knowmetrics.org/en/artifacts-2>

<https://rrchnm.org/what-we-do>

3. *SaaS* es el acrónimo de *Software as a Service*; en castellano se suele utilizar también el acrónimo *ScuS* – Software como un Servicio.

4. Cuota de mercado de los CMS de código abierto:

<https://www.opensourcems.com/cms-market-share>

5. A partir de la 7 incorpora en el *core* el módulo *Content Creation Kit (CCK)* para la creación de modelos de datos personalizados; a partir de la 8 incorpora el módulo *Views*, para la configuración asistida de vistas dinámicas en SQL.

6. Actualmente existen 47.211 módulos disponibles, de los cuales cerca de la mitad se encuentra en mantenimiento activo y es compatible con versiones superiores a la 7:

https://www.drupal.org/project/project_module

7. Por ejemplo, módulos cómo: *Features*, *Feeds*, *Teichi*, *Date YMD-Pre Epoch*, *Partial Date Values (Partial Date for Drupal)*, *Autocomplete Deluxe*, *Views data export*, etc.

8. En particular, la adopción del *framework Symfony*, del sistema de gestión de paquetes *PHP Composer* y del lenguaje de programación orientado a objetos, en lugar de la programación por procedimientos.

9. La versión 7 es utilizada por el 64% de todos los sitios web desarrollados en *Drupal*, mientras que la versión 8, después de seis años, ha alcanzado sólo el 29% del total:

<https://w3techs.com/technologies/details/cm-drupal>

10. *BackdropCMS* es usado por 1.706 webs:

<https://backdropcms.org/project/usage/backdrop>

11. Actualmente 533:

<https://backdropcms.org/modules>

aunque más de 70 módulos muy populares han sido incorporado en el *core*:

<https://backdropcms.org/upgrade-from-drupal/features-added-core>

y casi la totalidad de los 100 módulos más utilizados en *Drupal* ya han sido migrados:

<https://backdropcms.org/upgrade-from-drupal/top-100-drupal-7-modules>

12. Existen 1.380 distribuciones oficiales:

https://www.drupal.org/project/project_distribution

13. <https://us12.campaign-archive.com/?u=0b0b343dc81346ec97cfc51d9&id=acbdafa77c>

14. Los ejemplos de uso en HD proceden, por ejemplo, de la Lingüística (**Cimiano et al.**, 2020).

15. El motor de indexación y búsqueda Solr:

<https://lucene.apache.org/solr>

a pesar de ser presentado como una componente de base en la misma página del proyecto, sigue siendo un módulo opcional.

16. Otros casos de uso documentados son la cartografía del patrimonio en riesgo, la vigilancia de los efectos del aumento del nivel del mar en los recursos costeros, la investigación arqueológica, la gestión y publicación de información sobre la procedencia de los objetos de arte, etc. Una selección de proyectos que emplean *Arches* puede consultarse en:

<https://www.archesproject.org/implementations-of-arches>

17. <https://nodegoat.net/usecases>

18. <https://multilingualdh.org/en>

19. Las soluciones personalizadas, programadas desde cero, están menos sujetas, aunque no exentas, a los efectos negativos de la divulgación pública de los fallos de seguridad.

20. Más información al respecto puede encontrarse en:

<https://www.drupal.org/forum/newsletters/security-advisories-for-drupal-core/2014-10-15/sa-core-2014-005-drupal-core-sql>

y

<https://www.drupal.org/sa-core-2018-002>

21. Según los datos de *W3Techs*, PHP es usado por el 77,8% de todos los sitios web con un lenguaje de programación lado servidor conocido:

<https://w3techs.com/technologies/details/pl-php>

22. Por ejemplo, utilizando *Jekyll*, un generador de sitios web estáticos, en combinación con *Wax* o *CollectionBuilder*.

23. Los *flat-file* CMS, como *Grav*, son plataformas que no requieren base de datos y se basan en archivos de texto.

24. <https://www.archesproject.org/roadmap>

25. *LAMP* (*Linux*, *Apache*, *MySQL/MariaDB*, *PHP/Perl/Python*) es un acrónimo utilizado para definir un paquete de software muy utilizado para aplicaciones web. Cada letra representa uno de sus cuatro componentes de código abierto: *Linux* para el sistema operativo; *Apache* HTTP Server; *MySQL* para el sistema de gestión de bases de datos relacionales; *PHP* (o *Perl* o *Python*) para el lenguaje de programación.

8. References

Alcaraz-Martínez, Rubén (2012). "Omeka: exposicions virtuals i distribució de col·leccions digitals". *BiD: Textos universitaris de biblioteconomia i documentació*, v. 28.

<https://raco.cat/index.php/BiD/article/view/256877>

Alcaraz-Martínez, Rubén (2014). "CollectiveAccess, un sistema de gestión y difusión de colecciones de museos, archivos y bibliotecas". *BiD: Textos universitaris de biblioteconomia i documentació*, v. 33.

<https://doi.org/10.1344/BiD2014.33.23>

- Blanke, Tobias; Candela, Leonard; Hedges, Mark; Priddy, Mike; Simeoni, Fabio** (2010). "Deploying general-purpose virtual research environments for humanities research". *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, v. 368, n. 1925, pp. 3813-3828.
<https://doi.org/10.1098/rsta.2010.0167>
- Bocanegra-Barbecho, Lidia** (2020). "Ciencia ciudadana y memoria histórica: nuevas perspectivas historiográficas desde las humanidades digitales y la historia pública Digital". In: Caro, Jorge; Díaz-de-la-Fuente, Silvia; Ahedo, Virginia; Zurro, Débora; Madella, Marco; Galán, José-Manuel; Izquierdo, Luis R.; Santos, José-Ignacio; Del-Olmo, Ricardo. *Terra incognita: Libro blanco sobre transdisciplinariedad y nuevas formas de investigación en el Sistema Español de Ciencia y Tecnología*. PressBooks.
<https://doi.org/10.5281/zenodo.4034177>
- Bocanegra-Barbecho, Lidia; Toscano, Maurizio; Delgado-Anés, Lara** (2017). "Co-creación, participación y redes sociales para hacer historia. Ciencia con y para la sociedad". *Historia y comunicación social*, v. 22, n. 2, pp. 325-346.
<https://doi.org/10.5209/HICS.57847>
- Bree, Pim; Kessels, Geert** (2013). *Nodegoat: a web-based data management, network analysis & visualisation environment*.
<https://nodegoat.net>
- Brown, William J.; Malveau, Raphael C.; McCormick, Hays W.; Mowbray, Thomas J.** (1998). *AntiPatterns: refactoring software, architectures, and projects in crisis*. John Wiley & Sons, Ltd. ISBN: 978 0 471 19713 3
- Budin, Gerhard** (2015). "Digital humanities, language industry, and multilingualism: global networking and innovation in collaborative methods". In: *CIUTI-Forum 2014: Pooling academic excellence with entrepreneurship for new partnerships*. Peter Lang, pp. 423-448. ISBN: 978 3 034315708
- Burdick, Anne; Drucker, Johanna; Lunenfeld, Peter; Presner, Todd; Schnapp, Jeffrey** (2013). *Digital humanities*. Cambridge, MA: MIT Press. ISBN: 978 0 262528863
- Candela, Leonardo; Castelli, Donatella; Pagano, Pasquale** (2013). "Virtual research environments: an overview and a research agenda". *Data science journal*, v. 12, pp. GRDI75–GRDI81.
<https://doi.org/10.2481/dsj.GRDI-013>
- Care* (2020). *Corpus architecturae religiosae europeae (IV-X saec.)*.
<https://care.huma-num.fr/care/index.php>
- Carlisle, Philip K.; Avramides, Ioannis; Dalgity, Alison; Myers, David** (2014). "The Arches heritage inventory and management system: a standards-based approach to the management of cultural heritage information". *Cidoc (International Committee for Documentation of the International Council of Museums)*, v. 8.
<https://www.wmf.org/publication/arches-heritage-inventory-and-management-system-standards-based-approach-management>
- Causser, Tim; Wallace, Valerie** (2012). "Building a volunteer community: results and findings from transcribe bentham". *Digital humanities quarterly*, v. 6, n. 2.
<http://www.digitalhumanities.org/dhq/vol/6/2/000125/000125.html>
- Chevalier, Pascale; Granjon, Ludovic; Leclercq, Éric; Millereux, Arnaud; Savonnet, Marinette; Sapin, Christian** (2012). "Base de données annotées et wiki pour la constitution du corpus numérique CARE". *Hortus artium medievalium*, v. 18, n. 1, pp. 27-35.
<https://doi.org/10.1484/J.HAM.1.102782>
- Cimiano, Philipp; Chiarcos, Christian; McCrae, John P.; Gracia, Jorge** (2020) "Linguistic linked data in digital humanities". *Linguistic linked data*. Cham: Springer.
https://doi.org/10.1007/978-3-030-30225-2_13
- Davis, Edie; Heravi, Bahareh** (2021). "Linked data and cultural heritage: a systematic review of participation, collaboration, and motivation". *Journal on computing and cultural heritage*, v. 14.
<https://doi.org/10.1145/3429458>
- Dombrowski, Quinn** (2016a). "Drupal and other content management systems". In: Crompton, Constance; Lane, Richard; Siemens, Ray. *Doing digital humanities. Practice, training, research*. London: Routledge, pp. 289-302.
<https://doi.org/10.4324/9781315707860>
- Dombrowski, Quinn** (2016b). *Drupal for humanists (coding for humanists)*. Texas A&M University Press. ISBN: 978 1 623494728
- Dombrowski, Quinn** (2018). *Enterprise tools and DH*. Stanford | Digital humanities.
<https://digitalhumanities.stanford.edu/enterprise-tools-and-dh>

- Gallini, Stefania; Noiret, Serge** (2011). "La historia digital en la era del web 2.0. Introducción al dossier historia digital". *Historia crítica*, v. 43, pp. 16-37.
http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0121-16172011000100003
- Gilgado, José M.** (2014). *Avoiding the law of the instrument*. Josemdev.
<https://josemdev.com/avoiding-the-law-of-the-instrument>
- González-Pérez, César** (2018). *Information modelling for archaeology and anthropology. Software engineering principles for cultural heritage*. Springer.
<https://doi.org/10.1007/978-3-319-72652-6>
- Gruber, Ethan** (2010). *Building Omeka exhibits with Fedora repository content | Scholars' Lab*.
<https://scholarslab.lib.virginia.edu/blog/building-omeka-exhibits-with-fedora-repository-content>
- Hedges, Mark; Dunn, Stuart** (2018). *Academic crowdsourcing in the humanities*. Chandos Publishing. ISBN: 978 0 081009413
- Hirschheim, Rudy; Klein, Heinz** (2012). "A glorious and not-so-short history of the information systems field". *Journal of the association for information systems*, v. 13, n. 4.
<https://doi.org/10.17705/1jais.00294>
- Jiménez, Àngels** (2003). "Digital asset management: la gestión de la información multimedia en las organizaciones". *El profesional de la información*, v. 12, n. 6, pp. 452-461.
<http://profesionaldeinformacion.com/contenidos/2003/noviembre/4.pdf>
- Kaldeli, Eirini; Menis-Mastromichalakis, Orfeas; Bekiaris, Spyros; Ralli, Maria; Tzouvaras, Vassilis; Stamou, Giorgos** (2021). "CrowdHeritage: crowdsourcing for improving the quality of cultural heritage metadata". *Information*, v. 12, n. 2, 64.
<https://doi.org/10.3390/info12020064>
- Kräutli, Florian; Valleriani, Matteo** (2018). "CorpusTracer: A Cidoc database for tracing knowledge networks". *Digital scholarship in the humanities*, v. 33, n. 2, pp. 336-346.
<https://doi.org/10.1093/llc/fqx047>
- Lenz, Kirsten; Oberweis, Andreas** (1998). "Design of world wide web information systems". In: *Classification, data analysis, and data highways*, pp. 262-269. Berlin, Heidelberg: Springer.
https://doi.org/10.1007/978-3-642-72087-1_29
- Martínez-Caro, José-Manuel; Aledo-Hernández, Antonio-José; Guillén-Pérez, Antonio; Sánchez-Iborra, Ramón; Cano, María-Dolores** (2018). "A comparative study of web content management systems". *Information*, v. 9, n. 2, 27.
<https://doi.org/10.3390/info9020027>
- Nishanbaev, Ikrom; Champion, Erik; McMeekin, David** (2019). "A survey of geospatial semantic web for cultural heritage". *Heritage*, v. 2, n. 2, pp. 1471-1498.
<https://doi.org/10.3390/heritage2020093>
- Noiret, Serge** (2018). "Digital public history". In: Dean, David (ed.). *A companion to public history*. John Wiley & Sons, Ltd., pp. 111-124.
<https://doi.org/10.1002/9781118508930.ch7>
- Oldman, Dominic; Tanase, Diana** (2018). "Reshaping the knowledge graph by connecting researchers, data and practices in researchspace". In: Vrandečić, Denny; Bontcheva, Kalina; Suárez-Figueroa, Mari-Carmen; Presutti, Valentina; Celino, Irene; Sabou, Marta; Kaffee, Lucie-Aimée; Simperl, Elena. *The semantic web – ISWC 2018*. Springer, v. 11137, pp. 325-340.
https://doi.org/10.1007/978-3-030-00668-6_20
- Oomen, Johan; Aroyo, Lora** (2011). "Crowdsourcing in the cultural heritage domain: opportunities and challenges". *Proceedings of the 5th international conference on communities and technologies - C&T '11*, pp. 138-149.
<https://doi.org/10.1145/2103354.2103373>
- Pons, Anaclet** (2018). "El pasado fue analógico, el futuro es digital. Nuevas formas de escritura histórica". *Ayer*, v. 110, pp. 19-50.
https://revistaayer.com/sites/default/files/articulos/110-1-ayer110_HistDigital_APons_MEiroa.pdf
- Ridge, Mia** (2014). *Crowdsourcing our cultural heritage*. Ashgate: Routledge. ISBN: 978 1 472410221
- Rodríguez-Ortega, Nuria** (2018). "Cinco ejes para pensar las humanidades digitales como proyecto de un nuevo humanismo digital". *Artnodes*, v. 22.
<https://doi.org/10.7238/a.v0i22.3263>

- Ruest, Nick; Stapelfeldt, Kirsta** (2014). *Introduction to Islandora*. YorkSpace Repository.
<http://hdl.handle.net/10315/28006>
- Rutherford, Eleanor; Hemati, Wahed; Mehler, Alexander** (2018). "Corpus2Wiki: a mediawiki based annotation & visualisation tool for the digital humanities". In: Burghardt, Manuel; Müller-Birn, Claudia (eds.). *INF-DH-2018*. Bonn: Gesellschaft für Informatik e.V.
<https://doi.org/10.18420/INFDH2018-08>
- Sánchez-Nogales, Elena** (2019). "ComunidadBNE: el proyecto de enriquecimiento colaborativo de la Biblioteca Nacional de España". *Mi biblioteca: la revista del mundo bibliotecario*, v. 59, pp. 68-72.
- Scheuermann, Leif; Kroeze, Jan H. Hendrik** (2017). "Digital humanities and information systems: innovating two research traditions". *AMCIS 2017 Proceedings*. 3.
<https://aisel.aisnet.org/amcis2017/PhilosophyIS/Presentations/3>
- Scholz, Martin; Görz, Guenther** (2012). "WissKI: a virtual research environment for cultural heritage". *ECAI*, pp. 1017-1018.
<https://doi.org/10.3233/978-1-61499-098-7-1017>
- Sciotti, Elisa** (2019). "Il transcribathon: un nuovo approccio alle lettere manoscritte risalenti alla Grande Guerra". *Digitalia*, v. 2, pp. 116-122.
<http://digitalia.sbn.it/article/view/2174/1503>
- Simpson, Robert; Page, Kevin R.; De Roure, David** (2014). "Zooniverse: observing the world's largest citizen science platform". *Proceedings of the 23rd international conference on world wide web*, pp. 1049-1054.
<https://doi.org/10.1145/2567948.2579215>
- Sommerville, Ian** (2005). *Ingeniería del software*. Pearson Educación SA. ISBN: 84 7829 074 5
- Stapelfeldt, Kirsta; Moses, Donald** (2013). "Islandora and TEI: current and emerging applications/approaches". *Journal of the text encoding initiative*, v. 5.
<https://doi.org/10.4000/jtei.790>
- Terras, Melissa** (2016). "Crowdsourcing in the digital humanities". In: Schreibman, Susan; Siemens, Ray; Unsworth, John (eds.). *A new companion to digital humanities*, 2nd ed. Wiley-Blackwell, pp. 420-439. ISBN: 978 1 118 68059 9
- Toscano, Maurizio** (2018). "Where the researcher cannot get: open platforms to collaborate with citizens on cultural heritage research data". In: Romero-Frías, Esteban; Bocanegra-Barbecho, Lidia (eds.). *Ciencias sociales y humanidades digitales aplicadas*. Granada: Universidad de Granada, pp. 538-561.
<https://doi.org/10.5281/zenodo.1469337>
- Toscano, Maurizio; Rabadán, Aroa; Ros, Salvador; González-Blanco, Elena** (2020). "Digital humanities in Spain: historical perspective and current scenario". *Profesional de la información*, v. 29, n. 6.
<https://doi.org/10.3145/epi.2020.nov.01>
- Transcribe Bentham* (2021). *Transcription desk*.
http://transcribe-bentham.ucl.ac.uk/td/Transcribe_Bentham
- Velios, Athanasios; Martin, Aurelie** (2017). "Off-the-shelf CRM with *Drupal*: a case study of documenting decorated papers". *International journal on digital libraries*, v. 18, n. 4, pp. 321-331.
<https://doi.org/10.1007/s00799-016-0191-5>
- Whitten, Jeffrey L.; Bentley, Lonnie D.; Dittman, Kevin C.** (2004). *Systems analysis and design methods*. McGraw-Hill Irwin. ISBN: 978 0 071215213
- Wiberg, Andrew** (2014). "Murkuntu: information retrieval system engineered for indigenous individuals and communities". In: *IFLA WLIC 2014*, 16-22 August 2014, Lyon, France.
<http://library.ifla.org/id/eprint/922>