

MARAFd (Modelo de Archivo Racional Abierto de Fondo documental digital)

Historia y origen

MARAFd fue creado por **Rafael Lomeña Varo** (Alicante, España) como una solución práctica para gestionar colecciones digitales extensas. En 2022 inició el proyecto orientado al “Fondo Documental Digital Abierto sobre la ciudad de Melilla y la región del Rif” ¹ ², motivado por la fragmentación de documentos históricos en la red. El objetivo fundacional era ofrecer una plataforma libre y ligera de **preservación y divulgación** documental sobre Melilla-Rif, unificando contenido disperso y facilitando el acceso público ³. La iniciativa, de carácter personal (no institucional), pretende ser un “proyecto abierto de preservación documental digitalizada” con fines divulgativos ³.

Arquitectura técnica

Técnicamente, MARAFd es un *ecosistema de software de código abierto* basado en tecnologías web sencillas. Según la documentación oficial, “surge a partir de una integración estratégica de tecnologías abiertas y gratuitas como **PHP, Batch-Script, Bootstrap, JavaScript y HTML5**” ⁴. El backend de búsquedas y navegación está implementado con **PHP** (rutinas *finder.php* y *navega.php*), mientras que el frontend usa **HTML5/JavaScript** y framework **Bootstrap** para ofrecer una interfaz web responsive ⁵. Un script **Batch (.cmd)** local genera el índice (“tesauro”) desde la estructura de carpetas, y la sincronización con el servidor se hace por **SFTP** ⁵ ⁶.

MARAFd **no utiliza ninguna base de datos relacional externa**: en su lugar mantiene un archivo de texto que actúa como índice-tesauro. Este índice se genera automáticamente a partir del **nombre completo (ruta)** de cada fichero, codificando en él metadatos clave (año, tipo, título, etc.) ⁶. Al subirlo al servidor, se mantiene un espejo de todo el repositorio sin necesidad de backups adicionales ⁶. Las búsquedas en línea recorren este índice plano: el “buscador rápido” sólo rastrea las cadenas de palabras en el tesauro generado desde los nombres de archivo, **no busca dentro del contenido de los documentos** ⁷. Así, MARAFd prioriza la accesibilidad instantánea (incluso con terabytes de datos) usando sólo el sistema de archivos del hosting ⁸ ⁶.

Aplicaciones prácticas y casos de uso

MARAFd se ha aplicado ya en fondos documentales reales. El caso más destacado es el “**Fondo Documental Digital Melilla-Rif**”: una colección temática abierta sobre la historia de Melilla y el Rif, con más de 9.000 archivos multimedia (más de 100 GB de datos) indexados ⁹. En apenas dos años, este repositorio – catalogado en el Directorio de colecciones de Hispana – se ha convertido en “referente para la investigación y divulgación” de esa historia local ⁹. Otro ejemplo es un “Fondo Documental Digital sobre Protohistoria e Historia Antigua del Mediterráneo Occidental” (acceso restringido), que emplea MARAFd para gestionar un vasto conjunto de documentos de las primeras civilizaciones mediterráneas ¹⁰. En ambos casos MARAFd ha permitido ofrecer **acceso rápido y completo** a colecciones de miles de documentos, adaptando la interfaz al perfil del investigador sin requerir inversiones en infraestructura compleja ¹¹. Estos ejemplos ilustran usos en archivos históricos y

proyectos culturales, desde iniciativas personales hasta entes públicos que agregan sus recursos en plataformas como Hispana.

Comparación con otras tecnologías de gestión documental

Característica	MARAFd	DSpace	Archivematica	Alfresco
Tecnología base	PHP, HTML5, JavaScript, Batch (LAMP) ⁴ Sin BD externa ⁶	Java (Apache Tomcat), PostgreSQL (BD relacional) ¹²	Python (Django), bases de datos (e.g. PostgreSQL/MySQL) ¹³	Java EE (Spring), base de datos relacional ¹⁴
Modelo de datos	Índice basado en nombres de archivo (tesauro) ⁶ (no esquema formal)	Metadatos estructurados (p. ej. Dublin Core), OAI-PMH ¹⁵	Metadatos de preservación (METS, PREMIS, AIP) ¹³	Metadatos y contenidos empresariales (flexible, custom)
Función principal	Gestión de archivos digitales y fondo documental ligero; sin dependencia de software específico.	Repositorio institucional de contenidos académicos/publicaciones.	Sistema de preservación digital y workflows (ingestión a largo plazo).	Gestión integral de documentos/contenido empresarial (ECM).
Búsqueda	Solo sobre índice (palabras en tesauro) ⁷ ; sin indexación interna de archivos.	Búsqueda por metadatos y texto completo (Lucene/Solr) y OAI-PMH ¹⁵ .	Enfocado al flujo de ingestión; no ofrece portal público de búsqueda de contenidos.	Búsqueda de texto completo e indexación avanzada (incluye datos, imágenes, versiones).
Escalabilidad	Prácticamente ilimitada según hosting (mil GB hasta varios TB) ¹⁶ . Baja carga de sistema.	Escalable con recursos (depende de BD y servidor); usado en miles de repositorios.	Altamente escalable, pero orientado a procesamiento batch complejos.	Altamente escalable en entorno empresarial (cluster, cloud, BD distribuida).
Licencia / Soporte	Código abierto (CC BY-NC-SA); desarrollo individual.	Código abierto (BSD), amplia comunidad académica (Lyrisis, DuraSpace).	Código abierto (AGPL) ¹⁷ ; comunidad internacional (Artefactual).	Edición Community OSS (LGPL/CPL); versión Enterprise comercial.

En síntesis, MARAFd destaca por su **bajo peso y simplicidad técnica** (solo ~10 KB de código), independencia de DB y mínimo consumo de recursos ¹⁸ ⁶ . DSpace, Archivematica o Alfresco ofrecen entornos más completos (bases de datos relacionales, flujos de trabajo, interfaces avanzadas), pero a

costa de una infraestructura más compleja y costosa ¹² ¹⁴ ¹³. MARAFd, por el contrario, sacrifica algunas capacidades (como búsqueda interna o flujos multiusuario) a favor de la **fácil implementación y mantenimiento**.

Guía de implementación

Requisitos técnicos: Para el administrador de MARAFd solo se necesita **almacenamiento suficiente** local y en servidor (hosting-web) para todos los archivos del fondo ¹⁹. El servicio de hosting debe permitir **conexión SFTP** y ejecución de código **PHP** (necesario para *finder.php* y *navega.php*) ²⁰. No se requieren bases de datos adicionales. En el cliente (usuario final) basta con un ordenador con navegador e Internet ²¹.

Pasos de instalación:

1. **Preparar repositorio local:** organizar todos los documentos en carpetas según la clasificación deseada. Conviene usar nombres de archivo descriptivos (ej. incluir año, autor, título).
2. **Generar índice (tesauro):** ejecutar el script *Batch* (p. ej. `reindizar.cmd`) en el equipo local para escanear las carpetas y crear el archivo de tesauro (índice) ⁶. Este fichero plano contendrá la ruta completa y metadatos codificados.
3. **Sincronizar al servidor:** subir todos los documentos al servidor web vía SFTP. Copiar también el archivo índice generado al directorio público. En paralelo, siempre se mantiene un *mirror* del contenido local en el servidor ⁶.
4. **Deploy de interfaz:** subir las rutinas PHP (*finder.php* para búsqueda rápida, *navega.php* para navegación por carpetas) y la página HTML de acceso (*index.html*) al servidor ⁵. Editar estos archivos para apuntar al directorio raíz de los documentos según convenga. Puede usarse cualquier editor de texto y cliente FTP (VSCode, FileZilla, Notepad++, etc.) ²⁰.
5. **Probar y ajustar:** acceder a la web, probar la búsqueda y la navegación. Al añadir nuevos documentos en el repositorio, repetir el paso 2 y volver a subir el nuevo índice. No se requieren reinicios del sistema ni configuraciones avanzadas.

Mantenimiento y escalabilidad: El mantenimiento consiste básicamente en volver a generar el índice cada vez que cambien o se agreguen archivos. Dado que todo es plano y está “espejado” en el servidor, no se precisan copias de seguridad adicionales ⁶. MARAFd puede escalar a repositorios muy grandes: se han demostrado volúmenes de **centenares de gigabytes a terabytes** sin necesidad de hardware especial ¹⁶. El límite práctico es el espacio y ancho de banda del hosting contratado. En suma, MARAFd admite crecer en miles o millones de archivos con un mínimo esfuerzo administrativo.

Evaluación general (ventajas, limitaciones y futuro)

Entre las **ventajas** principales de MARAFd sobresalen su **ligereza, escalabilidad y bajo coste**. Con “apenas 10 Kilobytes de código” puede manejar fondos de cientos o miles de gigabytes ¹⁸. Al prescindir de bases de datos, simplifica la instalación y el mantenimiento, “manteniendo un mirror de todos los directorios sin necesidad de backups” adicionales ⁶. La solución prioriza la **accesibilidad inmediata** al 100% del contenido indexado ⁶ ¹⁶. Según la memoria técnica oficial, MARAFd ofrece “una forma segura... con una escalabilidad prácticamente ilimitada y unos costes ínfimos” ¹⁶ ²². Además, la interfaz es ágil y personalizable, con capacidad incluso para incorporar herramientas de inteligencia artificial (por ejemplo, consultas mediante modelos GPT) en fase experimental ²³ ²⁴.

Entre las **limitaciones** se encuentra precisamente la simplicidad de su modelo de indexación. El buscador integrado **no analiza el contenido interno** de los archivos, solo los nombres codificados en el tesauro ⁷. Esto significa que no hay búsqueda de texto completo en línea; para eso habría que

descargar los documentos o usar sistemas externos de indexación. La dependencia de nombres de fichero “descriptivos” exige disciplina al nombrar archivos ²⁵. Tampoco incluye un sistema de permisos o usuarios avanzado (es básicamente de un único administrador), ni flujos de trabajo complejos como los que ofrecen sistemas empresariales.

Oportunidades de mejora: Se podrían integrar buscadores internos (p.ej. motores de búsqueda adaptados) o indexación de contenido para consultas más profundas. También sería útil soportar formatos de metadatos estándar (METS, Dublin Core) y protocolos como OAI-PMH para interoperabilidad institucional. La adición de interfaces colaborativas y control de versiones ampliaría su aplicabilidad. Además, al ser código abierto bajo licencia CC BY-NC-SA ²⁶, el proyecto puede crecer con aportes comunitarios: el autor ha anunciado que irá publicando más documentación en el futuro, por lo que es esperable que surjan mejoras y nuevos casos de uso.

Futuro del proyecto: MARAFd se perfila como una alternativa atractiva para archivos y bibliotecas con recursos limitados, especialmente en entornos hispanohablantes donde ya empieza a reconocerse (por ejemplo, por Hispana). Su enfoque disruptivo –gestionar grandes colecciones digitales con herramientas mínimas– responde al reto del “crecimiento exponencial de datos” que la propia documentación del autor destaca ¹⁸ ¹⁶. Si mantiene respaldo institucional (p.ej. a través de Hispana) y se difunde en la comunidad archivística, podría evolucionar hacia un estándar informal para repositorios ligeros. En cualquier caso, la evaluación general reconoce a MARAFd como una solución **muy eficiente y accesible** para la preservación y consulta de archivos digitales, con potencial para ampliarse mediante nuevas tecnologías de búsqueda e inteligencia artificial ¹⁸ ²⁴.

¹ ² ³ HISPANA | Directorio nacional de recursos digitales > DIRECTORY OF COLLECTIONS > Fondo documental digital abierto sobre la ciudad de...

<https://hispana.mcu.es/en/comunidades/registro.do?id=6260>

⁴ ⁵ ⁶ ⁸ ⁹ ¹⁰ ¹¹ ¹⁶ ¹⁸ ¹⁹ ²⁰ ²¹ ²² ²³ ²⁵ ²⁶ calentamientoglobalacelerado.net

https://calentamientoglobalacelerado.net/ebiblio_melilla_riff/MIS-OBRA-PROYECTOS/2025_memoria-tecnica-resumen-MARAFd-HISPANA_signed_autor.pdf

⁷ MARAFd: Gestión Inteligente de Fondos Documentales

https://calentamientoglobalacelerado.net/fondo_documental_Melilla_Riff/busqueda_avanzada.html

¹² ¹⁵ DSpace: An Open Source Dynamic Digital Repository

<https://www.dlib.org/dlib/january03/smith/01smith.html>

¹³ ¹⁷ Archivematica: open-source digital preservation system

<https://www.archivematica.org/>

¹⁴ Alfresco Docs - Alfresco Community Edition

<https://docs.alfresco.com/content-services/community/>

²⁴ Fondo Documental Digital MELILLA-RIFF

https://calentamientoglobalacelerado.net/fondo_documental_Melilla_Riff/