

EVENTOS

OAWeek 2024

2024

e

Migrando a DSpace 7: retos y oportunidades desde la perspectiva de cuatro universidades españolas. Experiencia en la UC3M

Rebeca Marín del Campo
Ana M^a Poveda Poveda

 **crue** Universidades
Españolas
I+D+i
Red de Bibliotecas
REBIUN



Carta de presentación

Año

2006

Datos

- Principalmente publicaciones resultado de la investigación de la universidad, en particular la financiada con fondos públicos
 - 36992 documentos
 - Gestionado por el Servicio de Biblioteca. es uno de los pilares de la política de Open Science del Vicerrectorado de Política Científica, coordinada a través de UniOS (Unidad/Universidad Open Science).
-

Personal

- **Equipo de e-Archivo: 2 técnicas biblioteca + 1 informática**
Adjunto a la Dirección y Coordinación de apoyo a la investigación-Ciencia abierta
Dirección del Servicio de Biblioteca e Investigador principal
Contrato de 1 técnico de apoyo al personal técnico para desarrollar actividades científico-técnicas
Unidad de Sistemas del Servicio de Informática y Comunicaciones
- ARVO: Empresa subcontratada para llevar a cabo los trabajos de consultoría y servicios vinculados a la actualización del software.

Hoja de Ruta

Mar 22 Concesión María de Guzmán

Recopilación y análisis de información

Oct 22

Reuniones con la empresa adjudicataria el Servicio de Biblioteca y la Unidad de Sistemas del SI. Análisis de necesidades

May 23

Reestructuración de comunidades y colecciones
Meta datación: nuevo metadato de afiliación e incorporación de los vocabularios COAR

Jul 23

Se establece el calendario para la migración
Solicitud formal de los servidores virtuales a la unidad de Sistemas para la migración

Feb 24

Cierre de la v. 5.10 para el envío de Docs. Se deja la consulta activa

Mar 24

Salida a producción.

Resultados clave

Mejorar las capacidades tecnológicas del repositorio institucional de la UC3M (e-Archivo)

Incrementar el depósito de resultados de investigación en el repositorio

Reforzar el compromiso de los investigadores con la Ciencia en abierto, facilitando el depósito de sus publicaciones.

Legitimar un mandato de depósito institucional de todas las publicaciones de la UC3M como parte de la Política institucional de Open Science en la que está trabajando la Universidad.

Resultados clave

Archivo Abierto Institucional de la Universidad Carlos III de Madrid

e-Archivo, el Repositorio Institucional de la Universidad Carlos III, tiene como objetivos reunir, archivar, preservar y difundir la Investigación de la Universidad, en particular, la financiada con fondos públicos.

Para los conjuntos de datos, consulte e-cienciaDatos.

Si es usted personal docente-investigador de la universidad y está interesado en depositar sus publicaciones, puede consultar la guía de autoarchivo de documentos o enviarnos un mensaje.

Con la colaboración del Ministerio de Ciencia e Innovación y de la Fundación Española para la Ciencia y la Tecnología (FECYT).



Buscar en el repositorio ...

🔍 Buscar



Desafíos

Cumplir los plazos



Minimizar el tiempo de inactividad



GIFs diseñados por Freepik

Retos técnicos y soluciones

Proceso de migración de datos

1. Se añade a los registros existentes los metadatos de afiliación (dc.affiliation.xxx): Departamento, Instituto, Grupo Investigación, Área, Cátedra y Servicio
2. Se migran los metadatos para compatibilidad nativa OpenAIRE 4
3. Se modifica la estructura de las comunidades, subcomunidades y colecciones. Se reduce considerablemente el número de colecciones:
 - ✓ Mejor rendimiento de la infraestructura
 - ✓ Agilización en los procesos de trabajo
 - ✓ Se mantiene la recuperación organizativa a través de los metadatos de afiliación
4. Se incorporan los vocabularios COAR para los metadatos de: Access Rights (Derechos de acceso), Resource Types (Tipos de resultados de investigación) y Version (Versión del recurso)
 - ✓ Se revisan los registros con 2 o más tipos documentales para dejar el más representativo, se mapean a la colección/colecciones del resto
5. Se cierra el envío de nuevos ítems, mantenemos la consulta, se migra la BD

Retos técnicos y soluciones

Hardware


1. Nueva máquina virtual en la infraestructura propia de la Universidad, mantenimiento a cargo del Grupo de Sistemas del Servicio de Informática.
2. Redimensionamiento de la máquina antes y durante el proceso de migración
 - ✓ 8 CPUs / 24GB RAM
 - ✓ Reserva 4GB para SOLR (Motor de búsqueda)
 - ✓ Límite a 4 CPUs máximo para PM2 (Gestor de procesos)

Retos técnicos y soluciones

Seguridad

Si disponemos de un cortafuegos de aplicaciones Web (WAF) que filtra y monitoriza el tráfico HTTP en nuestra institución, debemos tener en cuenta que:

- El WAF recibe las peticiones en el puerto 443 (https) y las redirige al servidor real a través del puerto 80 (http) exclusivamente, es decir: el servidor WAF convierte las peticiones

`https://e-archivo.uc3m.es`  `http://<mimaquina>.uc3m.es`

- Para acceder a nuestra máquina vía ssh no utilizaremos el alias, nos conectaremos a `<mimaquina>.uc3m.es`

Retos técnicos y soluciones

Personalizaciones / Funcionalidades específicas



Adaptación del módulo desarrollado para versiones anteriores

Configuración del módulo: definir datos de la institución, colecciones a “recolectar” (correspondencia colección RePEc = colección Repositorio) y los metadatos que irán en el fichero .rdf

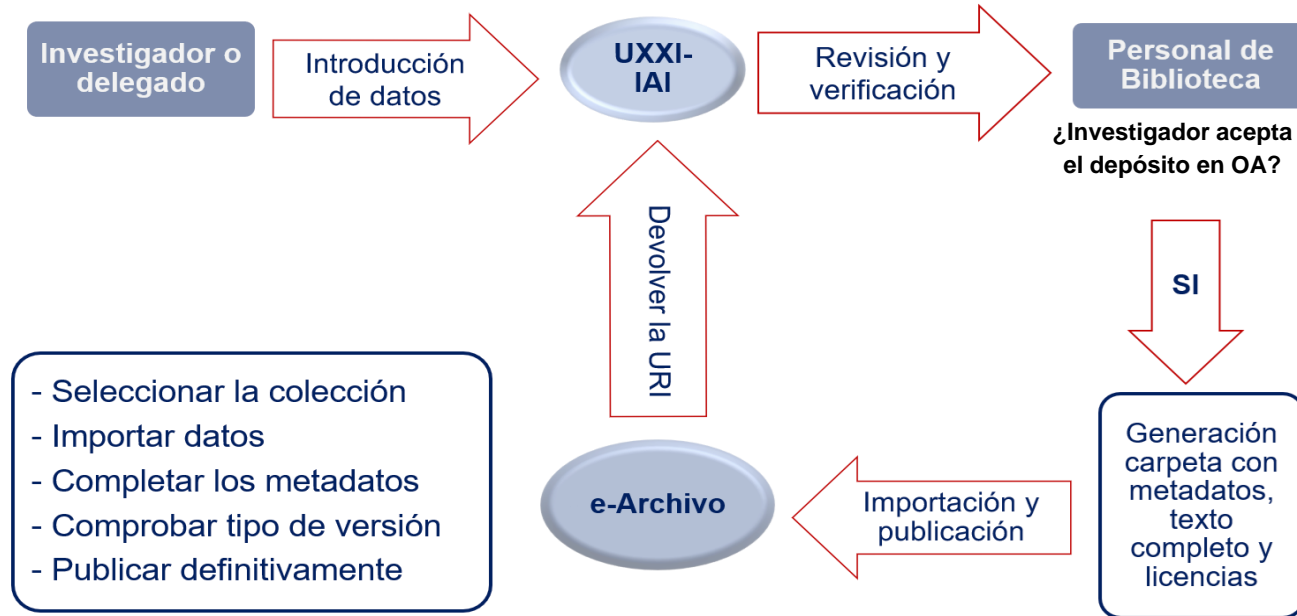
Prefijo = cte
Colecciones = werepe, derepe, dbrepe, ...
Tipo de colección = ReDIF-Paper
URL = <https://e-archivo.uc3m.es/>
Responsable = poveda@db.uc3m.es

Presenta los datos en un directorio virtual para la recolección

```
Mirroring cte at https://e-archivo.uc3m.es/repec/RePEc/cte/  
2024-10-10 01:56:30: Mirroring https://e-archivo.uc3m.es/repec/RePEc/cte/ with base https://e-archivo.uc3m.es/repec/RePEc/cte/ to E:\repec\remo\cte\  
2024-10-10 01:56:32: https://e-archivo.uc3m.es/repec/RePEc/cte/ => index.html  
2024-10-10 01:56:33: https://e-archivo.uc3m.es/repec/RePEc/cte/werepe/ => werepe\index.html  
2024-10-10 01:56:34: https://e-archivo.uc3m.es/repec/RePEc/cte/derepe/ => derepe\index.html  
2024-10-10 01:56:36: https://e-archivo.uc3m.es/repec/RePEc/cte/dbrepe/ => dbrepe\index.html  
2024-10-10 01:56:37: https://e-archivo.uc3m.es/repec/RePEc/cte/wbrepe/ => wbrepe\index.html  
2024-10-10 01:56:38: https://e-archivo.uc3m.es/repec/RePEc/cte/dsrepe/ => dsrepe\index.html  
2024-10-10 01:56:39: https://e-archivo.uc3m.es/repec/RePEc/cte/wsrepe/ => wsrepe\index.html  
2024-10-10 01:56:41: https://e-archivo.uc3m.es/repec/RePEc/cte/whrepe/ => whrepe\index.html  
2024-10-10 01:56:42: https://e-archivo.uc3m.es/repec/RePEc/cte/idrepe/ => idrepe\index.html  
2024-10-10 01:56:43: https://e-archivo.uc3m.es/repec/RePEc/cte/imrepe/ => imrepe\index.html  
2024-10-10 02:22:25: 1435 not modified documents checked  
2024-10-10 02:22:25: https://e-archivo.uc3m.es/repec/RePEc/cte/wsrepe/31230.rdf => wsrepe\31230.rdf  
2024-10-10 02:36:45: 801 not modified documents checked  
2024-10-10 02:36:45: 11 documents retrieved  
2024-10-10 02:36:45: 2236 unmodified documents  
2024-10-10 02:36:45: Mirror end
```


Retos técnicos y soluciones

Integración CRIS (Universitas XXI) / e-Archivo



Retos técnicos y soluciones

CRIS (Universitas XXI)

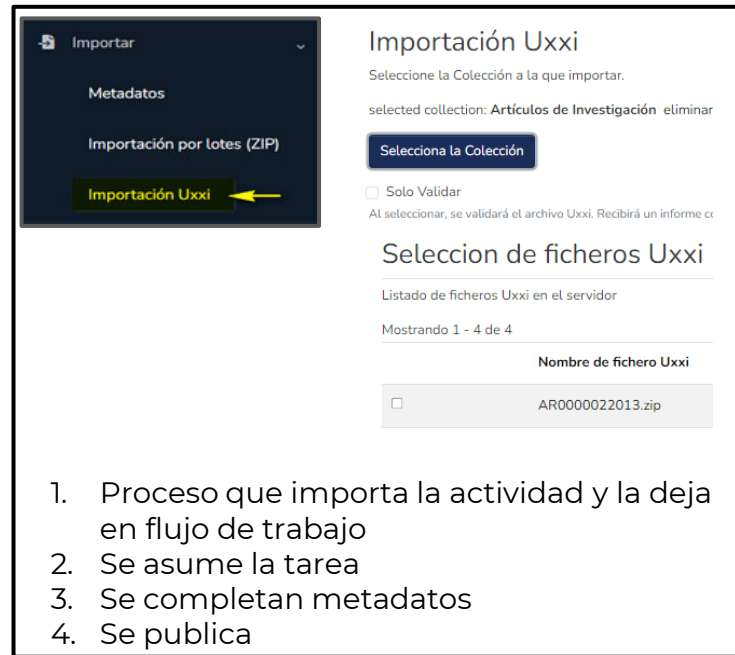
Pestaña e-Archivo:

- Investigador
 - Aceptar licencia depósito
 - Aceptar licencias CC
 - Subir texto completo (PDF)
- Biblioteca
 - Generar ZIP (carpeta con metadatos, texto completo y licencias)
 - Envío FTP entre máquinas

Una vez publicado:

- Webservice en UXXI que solicita a e-Archivo, las URI's de todas las actividades que no la tienen almacenada

e-Archivo



Importación Uxxi

Seleccione la Colección a la que importar.

selected collection: **Artículos de Investigación** eliminar

Solo Validar
Al seleccionar, se validará el archivo Uxxi. Recibirá un informe c

Selección de ficheros Uxxi

Listado de ficheros Uxxi en el servidor

Mostrando 1 - 4 de 4

Nombre de fichero Uxxi	
<input type="checkbox"/>	AR0000022013.zip

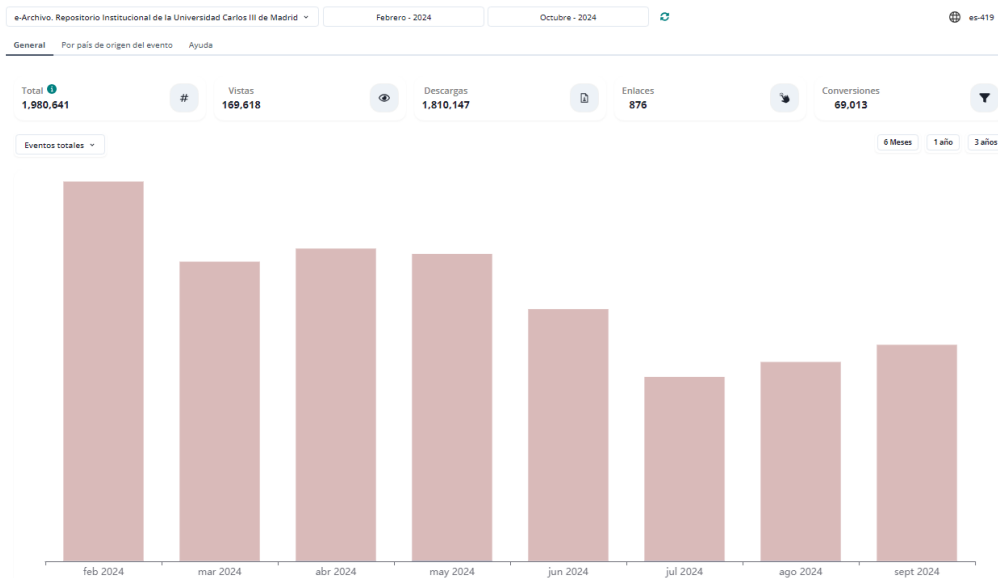
1. Proceso que importa la actividad y la deja en flujo de trabajo
2. Se asume la tarea
3. Se completan metadatos
4. Se publica

Retos para las bibliotecas

Módulo de estadísticas LA Referencia/Recolecta

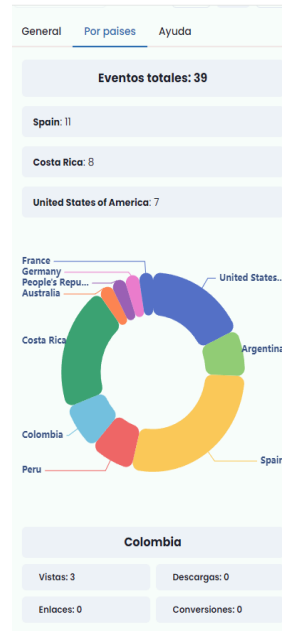
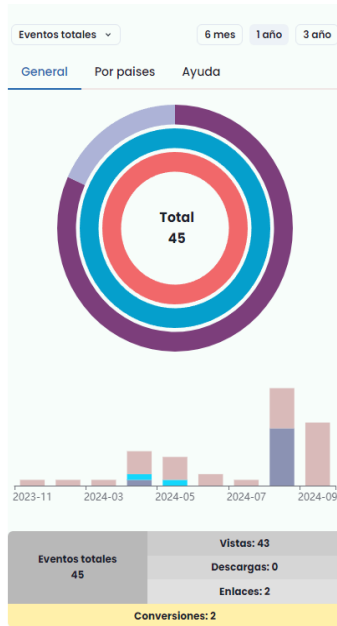
Recolector de datos estadísticos (dspace-stats-collector)

- Se instala de nuevo el SW del collector
- Se pone en marcha el envío, con fecha inicio el día del arranque de DSpace 7
- Se avisa a FECYT del cambio
- Se comprueba datos en el portal de RECOLECTA



Retos para las bibliotecas

Widget para datos del ítem (Pendiente instalación nueva versión)



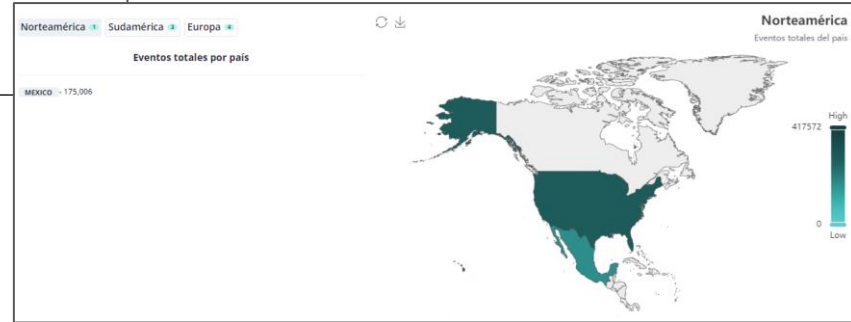
Retos técnicos y soluciones

Widget para datos globales del repositorio (el mismo que encontramos en el dashboard de RECOLECTA)



PAIS	VISTAS	DESCARGAS	ENLACES	CONVERSIONES
United States of America	16,102	278,600	107	4,743
Spain	58,856	358,641	75	27,896
Germany	2,409	56,380	0	658
Ukraine	410	102,998	0	41
Russian Federation	344	101,933	0	18
Mexico	16,920	158,005	81	5,890
Colombia	10,504	92,898	66	4,130
Peru	10,429	77,948	269	5,082
France	1,102	35,222	0	391
Argentina	5,012	55,417	37	2,071

[Descargar CSV](#)



Retos técnicos y soluciones

Script para envío masivo de datos retrospectivos (desarrollo disponible)

Enlaces de interés de LA Referencia

Documentación del proyecto: <https://ustats.lareferencia.info/>

Demo del widget de ítem: <https://widget-demo.netlify.app/>

Demo del widget del repositorio: <https://stats.lareferencia.info/demo/>