

# EVENTOS

XIX Workshop de Proyectos Digitales

2022

# e

**AVANCES EN EL USO DE UN SISTEMA DE ANÁLISIS ESTADÍSTICO EN REPOSITORIOS**

**Acción 3 “Métricas y estadísticas en Repositorios”**

**Grupo de Trabajo de Repositorios de REBIUN**

## MIEMBROS DEL GRUPO

**Ana Poveda** (UC3M) Coordinadora

**Ainara Cisneros** (UAH)

**Laura Frías** (UNICAN)

**Paco Martínez** (UPV)

**Toni Prieto** (UPC)



**Eva Navarro** (UPO)

**Sebastiano Giorgi** (UOC)

## OBJETIVO 1

### Sumar nuevos repositorios participantes en el envío de datos estadísticos a OpenAIRE

Indicadores:

- Conseguir, al menos, 5 repositorios más: en standby por el momento 
- Actualizar la información de la página web de REBIUN para utilizarla como canal de comunicación 
- Colaboración con OpenAIRE: propuestas de mejora, datos retrospectivos, ...

# EVENTOS

XIX Workshop REBIUN de Proyectos Digitales

2022

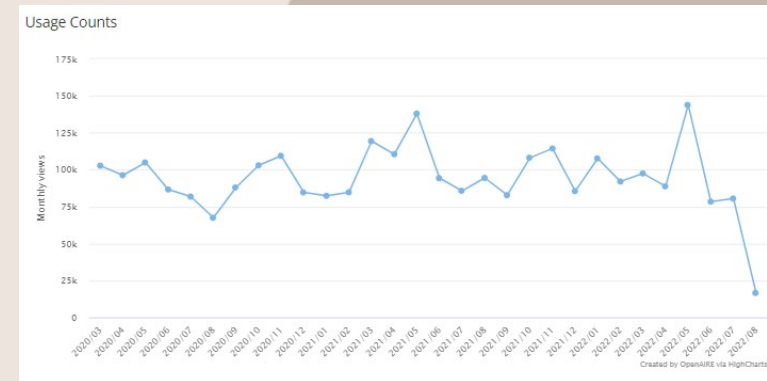
AVANCES EN EL USO DE UN SISTEMA DE ANÁLISIS ESTADÍSTICO EN REPOSITORIOS

Acción 3 “Métricas y estadísticas en Repositorios”

GT Repositorios REBIUN

## OpenAIRE

- Mayo 2022 → 14 repositorios universitarios enviando datos
- Junio 2022 → [Webinar](#) MONITOR – Institutional Dashboard
  - ❖ Datos sobre proyectos, financiadores, resultados ...
  - ❖ Servicio bajo demanda
  - ❖ [Monitoring Dashboard](#)
- Agosto 2022 → última actualización de datos
- Informes COUNTER 5 → pendiente, solo versión 4



1. Unique_Item_Investigations	• View Abstract
2. Total_Item_Investigations	• View Article preview
3. Unique_Item_Requests	• View HTML Full Text
4. Total_Item_Requests	• View PDF
	• Download PDF

# EVENTOS

XIX Workshop REBIUN de Proyectos Digitales

2022

AVANCES EN EL USO DE UN SISTEMA DE ANÁLISIS ESTADÍSTICO EN REPOSITORIOS

Acción 3 “Métricas y estadísticas en Repositorios”

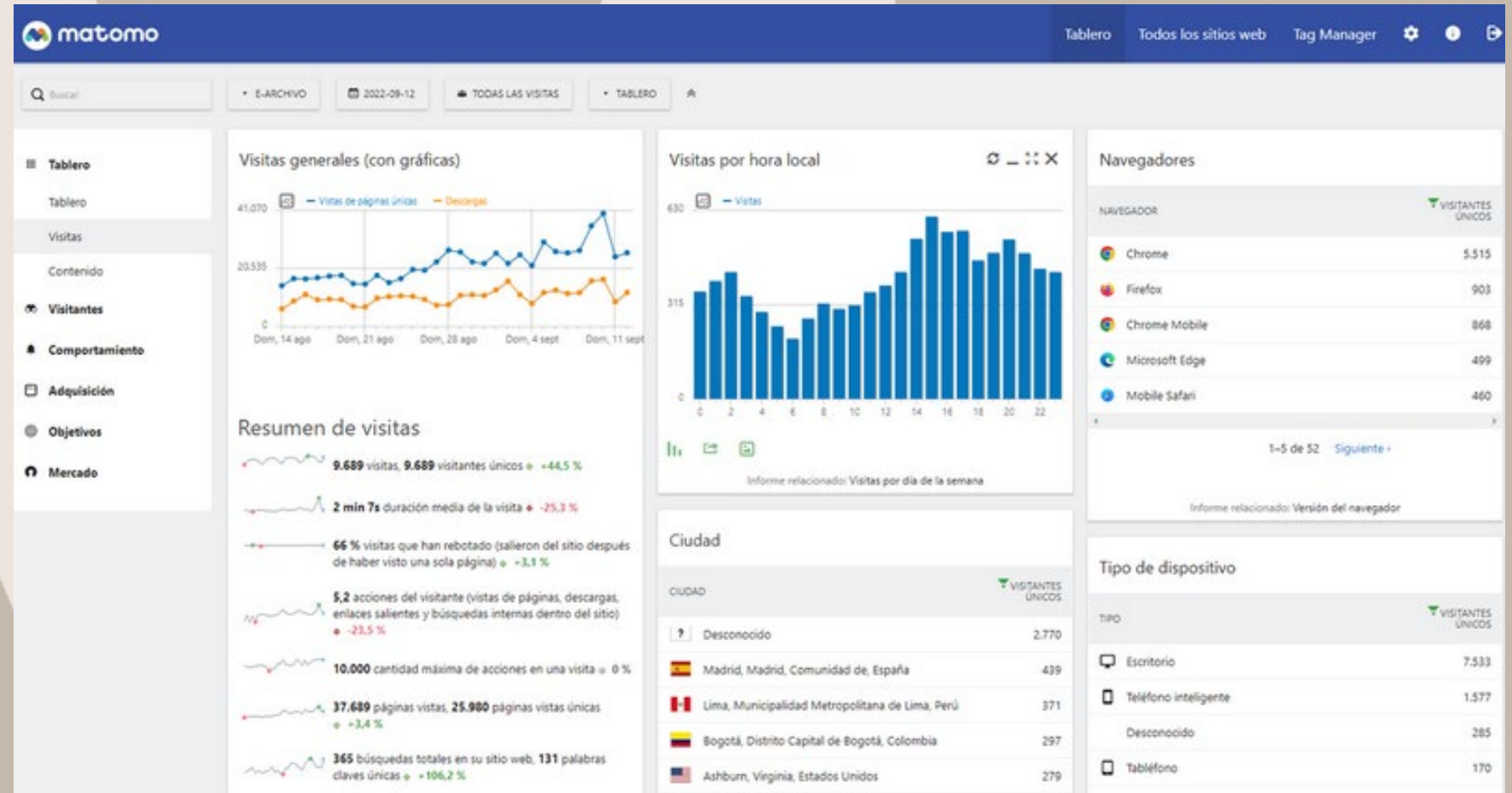
GT Repositorios REBIUN

## OBJETIVO 2

Avanzar en el uso de Matomo con una instalación propia

Indicador:

- Elaborar manual de uso de la consola



### OBJETIVO 3

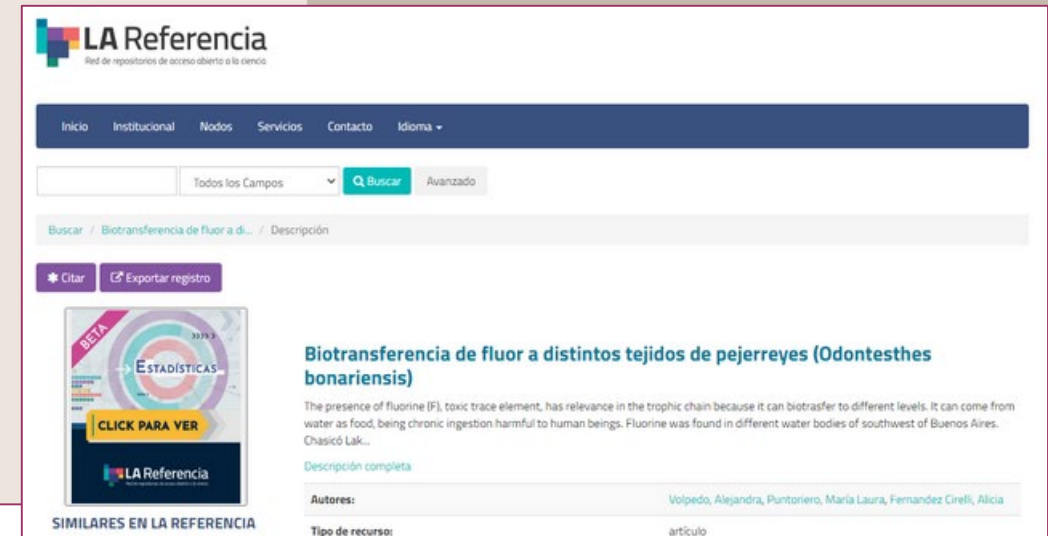
**Colaborar con FECYT en la puesta en marcha de Matomo como sistemas de estadísticas a nivel nacional**

Indicadores:

- Implementación del sistema de OpenAIRE en Recolecta → Descartado ❌
- Estudio del sistema de La Referencia ✓

## La Referencia

- **DSpace-stats-collector**: SW para el envío de datos, no incluido en el aplicativo DSpace
- **Matomo**: base de datos para capturar, normalizar y anonimizar los eventos de uso de cada repositorio
- **Amazon Glue**: preservación de eventos en Amazon S3, en formato Parquet (*orientado-a-columnas*)
- **Spark Jobs**: enriquecimiento con metadatos del Objeto Digital
- **Elasticsearch**: indexación de eventos
- **KIBANA**: monitorización y revisión de consistencia de datos
- **VEGA**: librería para visualización de las estadísticas



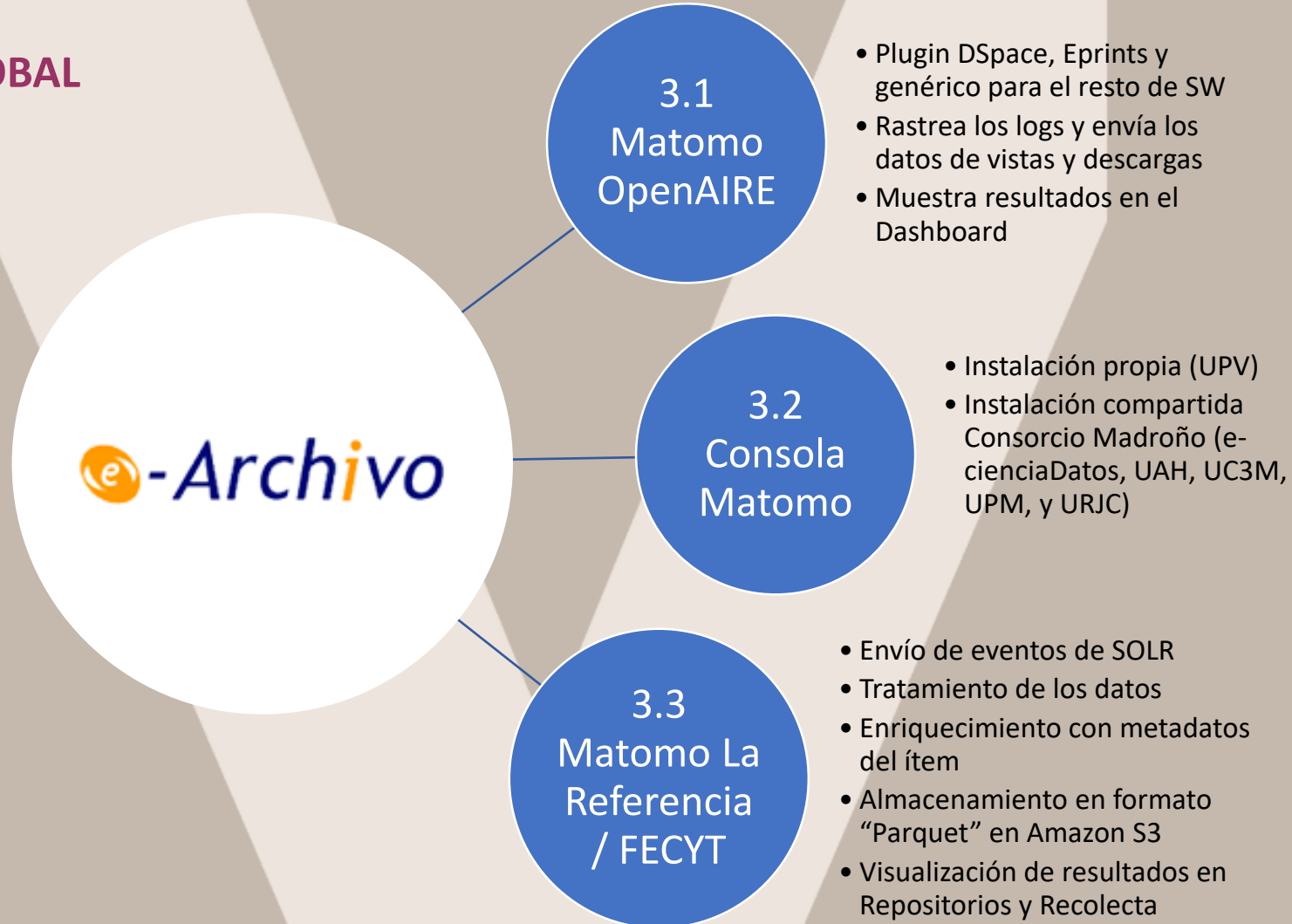
## PASOS

1. Reunión La Referencia / FECYT → Acuerdo de colaboración con 3 líneas de trabajo:
  - 1) Incluir DSpace-stats-collector en los repositorios españoles
  - 2) Incluir DSpace-stats-collector en Recolecta
  - 3) Visualización de estadísticas en Recolecta
2. Presentación del proyecto en la Jornada 15 de años de Recolecta
3. Reunión La Referencia / FECYT / Grupo Acción 3 GT REBIUN
  - 1) Grupo piloto de Repositorios DSpace → UAH, UC3M, UNICAN, UPC y UPV
  - 2) Caso especial repositorios con Oracle → Se ofrece la UPV para hacer las pruebas
  - 3) Posibilidad de desarrollo para Invenio → Se trabajará con la UAB
  - 4) Repositorios EPrints → pruebas con el genérico de OpenAIRE (La Referencia)
  - 5) Dspace 7 y DspaceCRIS → adaptación del DSpace-stats-collector (La Referencia )

## SITUACIÓN ACTUAL

1. Línea 1: instalar DSpace-stats-collector en los repositorios del piloto
  - 1) UPC: instalado y enviando datos a La Referencia
  - 2) UPV: instalado con la librería de Oracle funcionando, problemas de envío por volumen de datos
  - 3) UC3M: instalado, problemas de envío por volumen de datos
  - 4) UNICAN y UAH: pendientes de instalar
2. Línea 2: DSpace-stats-collector instalado en Recolecta (FECYT) ✓
3. Línea 3: Visualización de estadísticas en Recolecta → trabajando en su implementación (FECYT)
4. Invenio: Reunión pendiente con la UAB y La Referencia, en octubre

## VISIÓN GLOBAL



- Plugin DSpace, Eprints y genérico para el resto de SW
- Rastrea los logs y envía los datos de vistas y descargas
- Muestra resultados en el Dashboard

- Instalación propia (UPV)
- Instalación compartida Consorcio Madroño (e-cienciaDatos, UAH, UC3M, UPM, y URJC)

- Envío de eventos de SOLR
- Tratamiento de los datos
- Enriquecimiento con metadatos del ítem
- Almacenamiento en formato “Parquet” en Amazon S3
- Visualización de resultados en Repositorios y Recolecta

## MATOMO ANALYTICS



Matomo es un sistema de análisis estadístico similar a Google Analytics, con las siguientes características:

- ❖ Software Libre
- ❖ Interfaz fácil de usar
- ❖ Multilingüe
- ❖ Instalación propia o en la nube
- ❖ Facilita y prioriza la privacidad (mantiene el 100% de la propiedad de los datos)
- ❖ Se basa en PHP/MySQL
- ❖ Última versión 4.11
- ❖ Posibilidad de importar datos históricos de Universal Analytics

Más información en <https://matomo.org/>

**EVENTOS**

XIX Workshop REBIUN de Proyectos Digitales

**2022**

AVANCES EN EL USO DE UN SISTEMA DE ANÁLISIS ESTADÍSTICO EN REPOSITORIOS

Acción 3 “Métricas y estadísticas en Repositorios”

GT Repositorios REBIUN

**¡GRACIAS!**

