

FECYT Innovación

Métricas y Estadísticas en Repositorios

Grupo de Trabajo de Repositorios REBIUN



17 de mayo de 2022

Ana Poveda
anamaria.poveda@uc3m.es

Métricas y Estadísticas en Repositorios

Un poco de contexto

- **2006** aparecen los primeros repositorios españoles
- **2007** nacimiento de RECOLECTA (FECYT + REBIUN)
- Intentos de utilización de algún sistema de estadísticas por parte de RECOLECTA
- Matomo (Piwik en sus inicios)
 - OpenAIRE
 - **2018** Consorcio Madroño
- **Enero 2019** UC3M propone crear una acción, dentro del GT de Repositorios de REBIUN, sobre métricas y estadísticas

Métricas y Estadísticas en Repositorios

Grupo de Trabajo de Repositorios REBIUN

Miembros:

2019 Grupo inicial

Ainara Cisneros (Universidad de Alcalá de Henares, Madrid)

Laura Frías (Universidad de Cantabria)

Paco Martínez (Universitat Politècnica de València)

Ana Poveda, Coordinadora (Universidad Carlos III de Madrid)

2020

Toni Prieto (Universitat Politècnica de Catalunya)

2022

Sebastiano Giorgi (Universitat Oberta de Catalunya)

Eva Navarro (Universidad Pablo de Olavide de Sevilla)

Métricas y Estadísticas en Repositorios

Grupo de Trabajo de Repositorios REBIUN

Objetivos:

1. Monitorizar los avances de métricas y estadísticas en el ámbito de OpenAIRE como posible modelo a seguir en repositorios
2. Avanzar en el servicio de métricas y estadísticas
3. Colaborar con la FECYT
 1. Incluir una recomendación de uso en la guía de evaluación de repositorios (Objetivo cumplido)
 2. Puesta en marcha de Matomo como sistema de estadísticas a nivel nacional
4. Realizar tareas de difusión de nuestros avances en diferentes eventos

Métricas y Estadísticas en Repositorios

Matomo Analytics

¿Qué es Matomo Analytics?

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

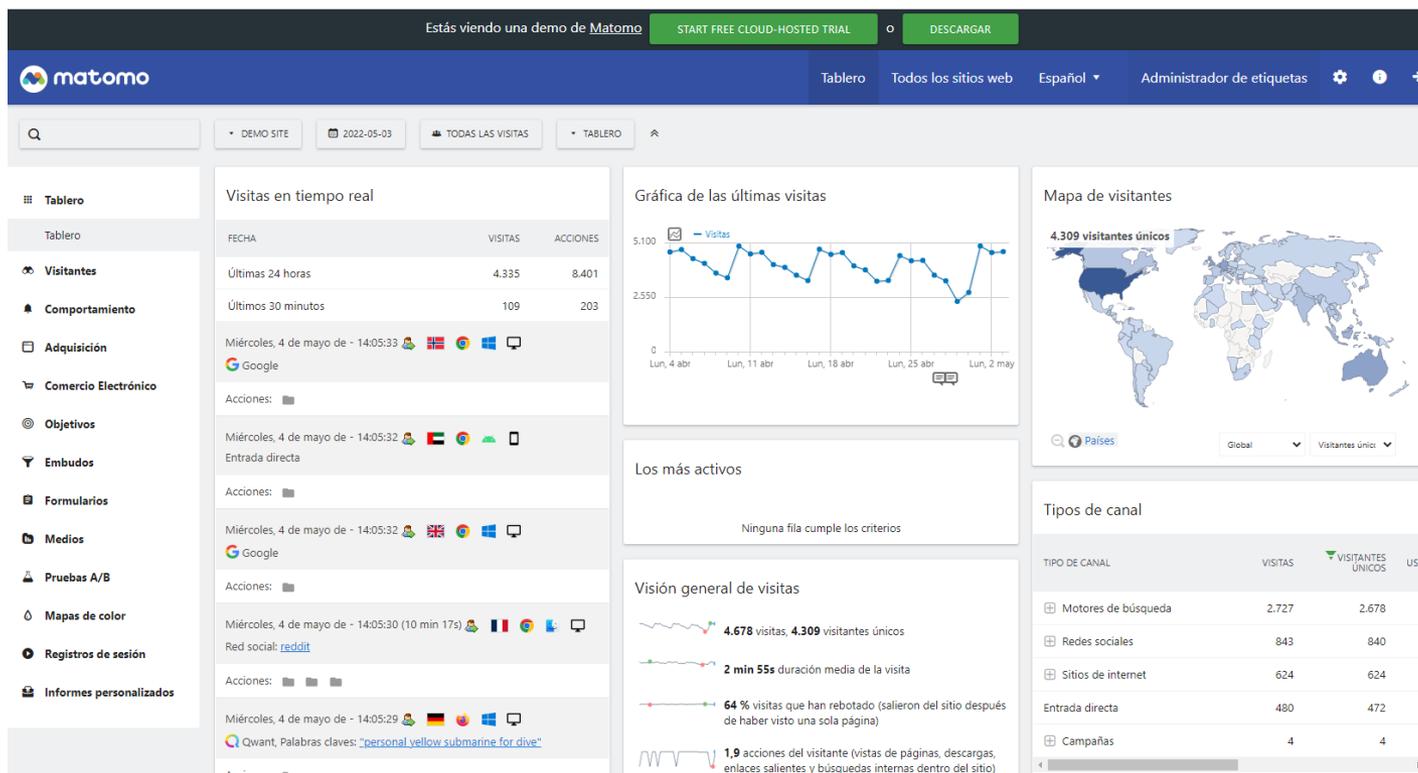
- ❖ Software Libre
- ❖ Interfaz fácil de usar
- ❖ Multilingüe
- ❖ Instalación propia o en la nube
- ❖ Facilita y prioriza la privacidad (mantiene el 100% de la propiedad de los datos)
- ❖ Se basa en PHP/MySQL
- ❖ Última versión 4.9
- ❖ Posibilidad de importar datos históricos de Universal Analytics (Google Analytics 4 no lo ofrece)

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

Métricas y Estadísticas en Repositorios

Matomo Analytics

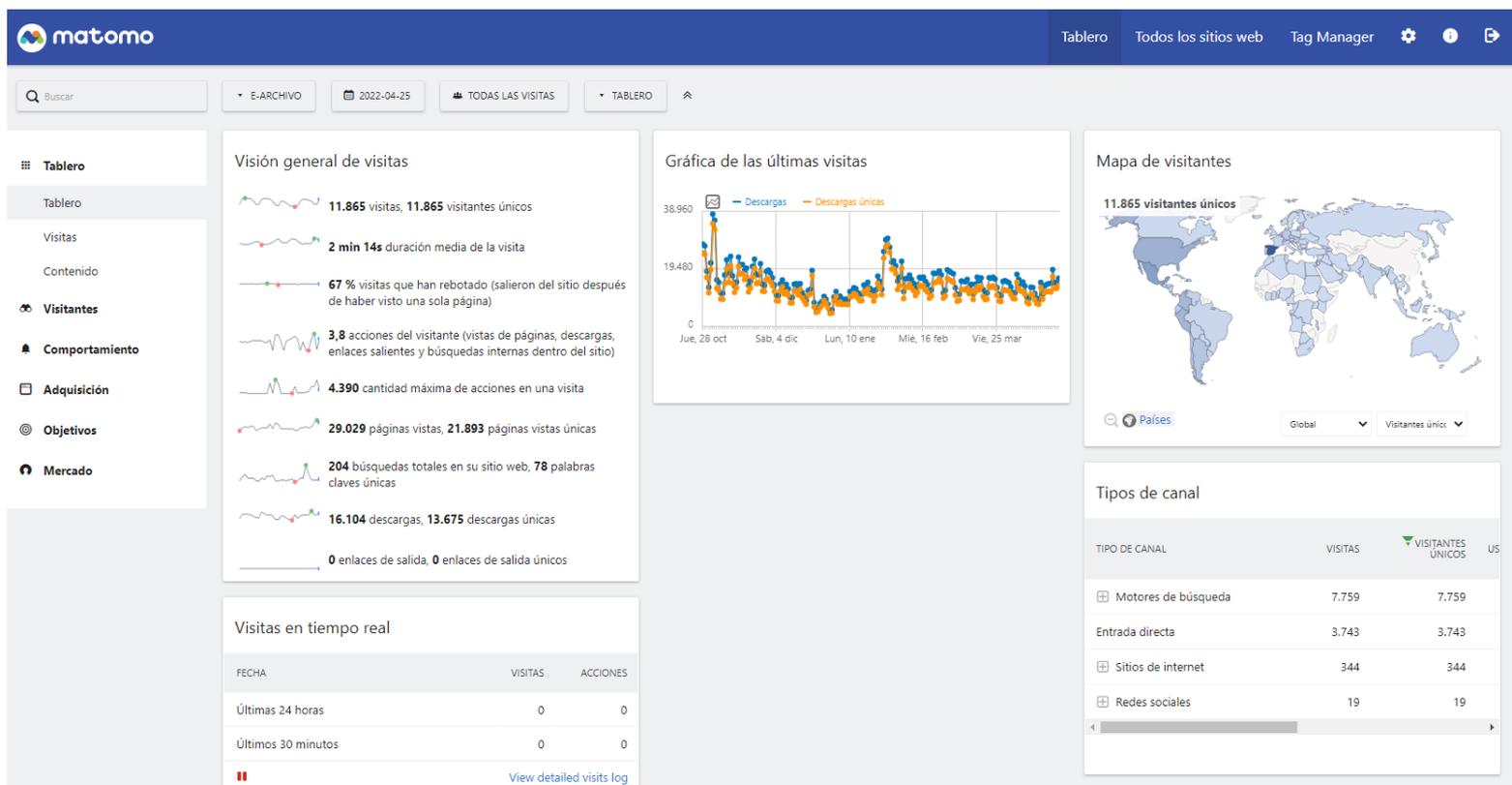
Demo de Matomo (Instalación en la nube)



Métricas y Estadísticas en Repositorios

Matomo Analytics

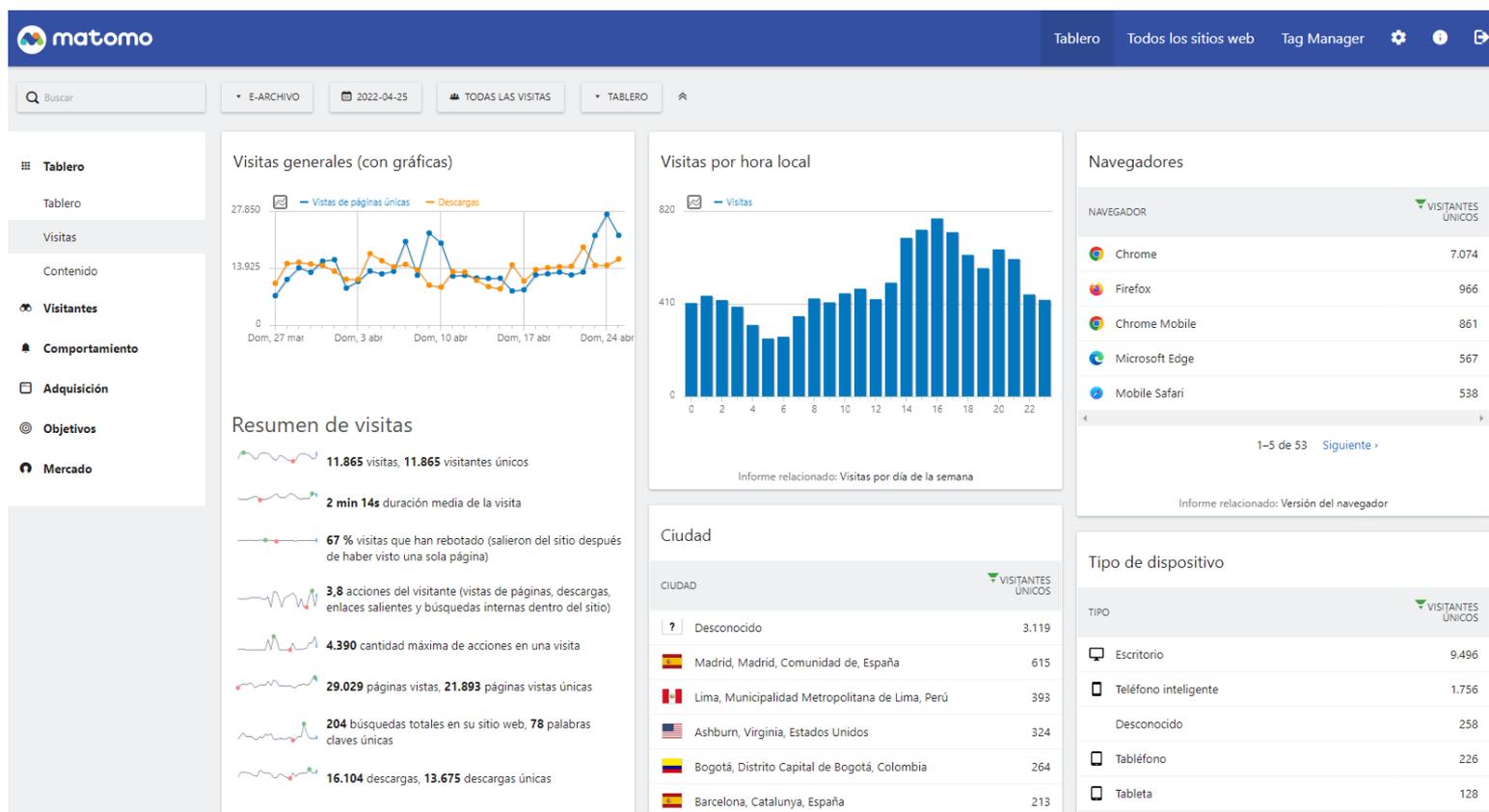
Matomo Consorcio Madroño (Tablero)



Métricas y Estadísticas en Repositorios

Matomo Analytics

Matomo Consorcio Madroño (Visitas)



Métricas y Estadísticas en Repositorios

Matomo Analytics en OpenAIRE

Arquitectura y flujo de trabajo

OpenAIRE opta por el uso de Matomo Analytics para la recogida y análisis de los datos de los repositorios registrados en su sistema

Coexisten dos sistemas de envío de datos estadísticos:

1. **PUSH:** Cada repositorio envía directamente sus datos a través del plugin desarrollado por OpenAIRE

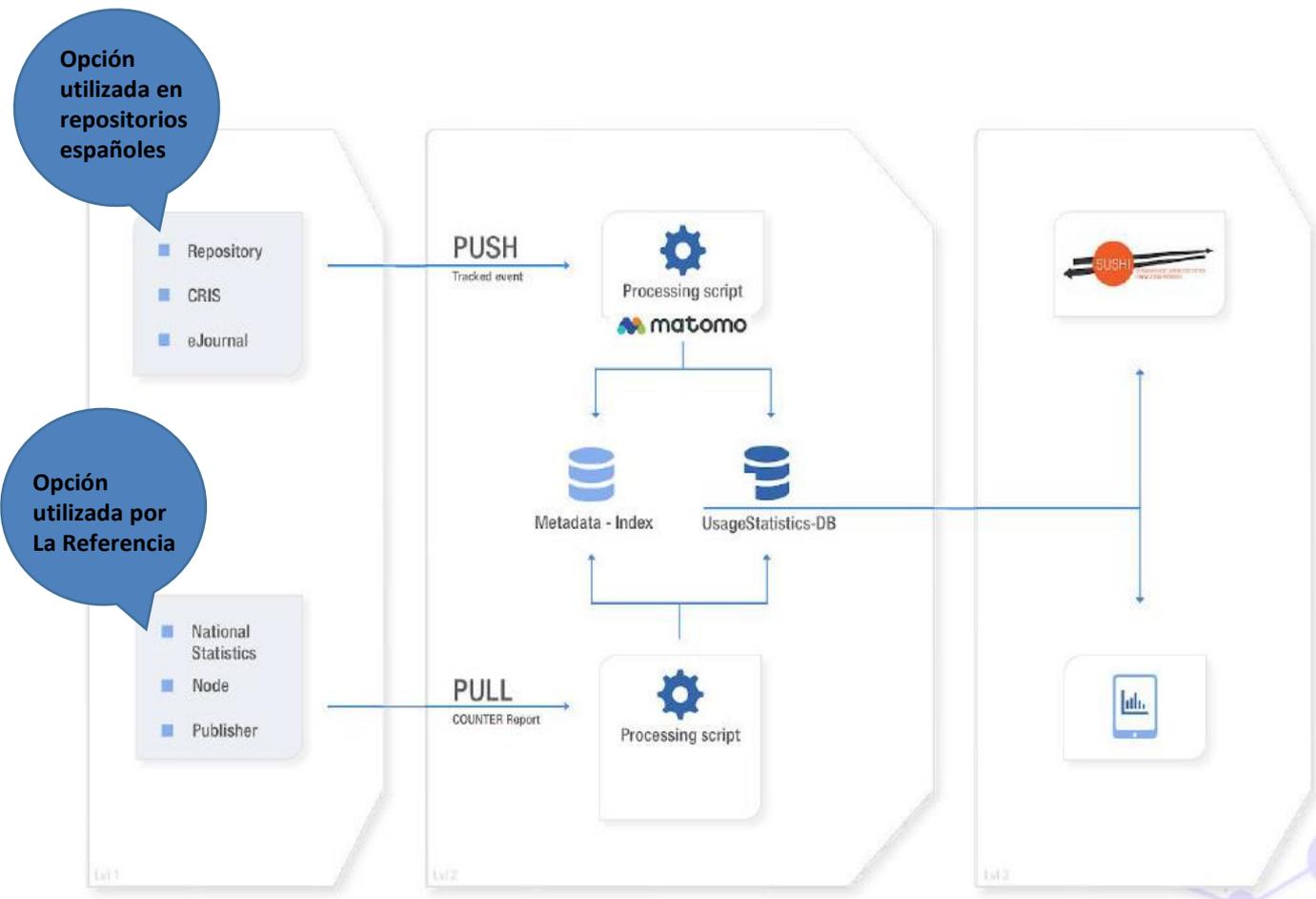
2022-03-30 12:30:54,071 INFO org.dspace.usage.LoggerUsageEventListener @
anonymous:session_id=F049C39FFCDB659C4F1D554D9AF22FCD:ip_addr=139.47.1.1:view_bitstream:bitstream_id=60445

2022-03-30 12:30:54,261 INFO com.openaire.piwik.tracker.OpenAirePiwikTracker @ Sent usage event to piwik

2. **PULL:** Un recolector nacional/regional (por ejemplo, La Referencia) recoge los datos estadísticos de los repositorios que cosecha y realiza el envío a OpenAIRE

Métricas y Estadísticas en Repositorios

Matomo Analytics en OpenAIRE



Métricas y Estadísticas en Repositorios

Matomo Analytics en OpenAIRE

Pasos para unirse a Matomo OpenAIRE

1. Registrar nuestro repositorio en OpenAIRE
2. Descargar el código de rastreo. Existen tres versiones:
 1. Dspace (<https://github.com/openaire/OpenAIRE-Piwik-Dspace>)
 2. EPrints (<https://github.com/openaire/EPrints-OAPiwik>)
 3. Genérico en Python (<https://github.com/openaire/Generic-Matomo-Tracker>)
3. Instalar y configurar el código en el repositorio añadiendo dos parámetros propios
 1. Matomoid
 2. Authenticationtoken
4. Validar con OpenAIRE que reciben nuestros datos

5. Aplicación de las reglas COUNTER sobre los datos de uso y almacenamiento en la base de datos
6. Procesamiento de eventos de uso y generación de informes estadísticos (COUNTER V.4)
7. Presentación de los resultados en el Dashboard del repositorio en OpenAIRE

R
E
P
O
S
I
T
O
R
I
O

OpenAIRE

Métricas y Estadísticas en Repositorios

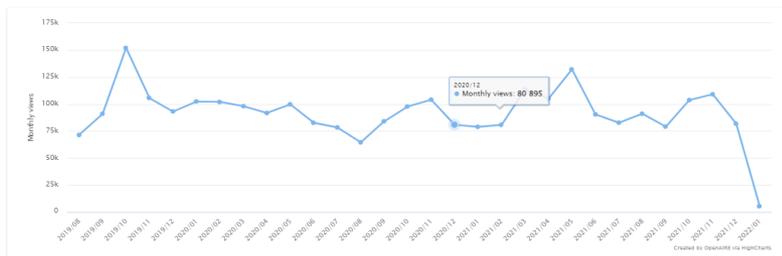
Dashboard OpenAIRE

RESUMEN Y GRÁFICOS

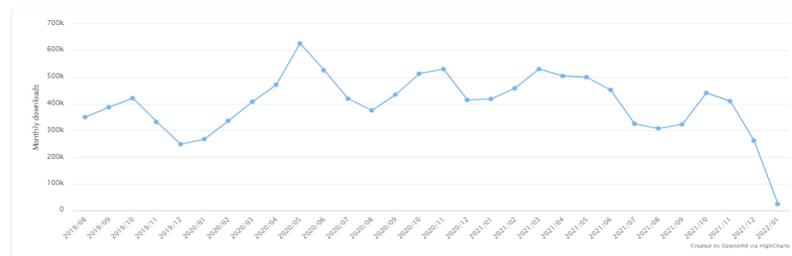


1. Vistas de registros en el portal de OpenAIRE
2. Vistas de registros en el repositorio
3. Descargas de documentos realizadas desde el repositorio

Totales



Vistas mensuales



Descargas mensuales

Métricas y Estadísticas en Repositorios

Dashboard OpenAIRE

INFORMES

DASHBOARD UPDATE AGGREGATION HISTORY ENRICHMENTS USAGE COUNTS

Get usage counts report for Repositorio Institucional de la Universidad Carlos III de Madrid

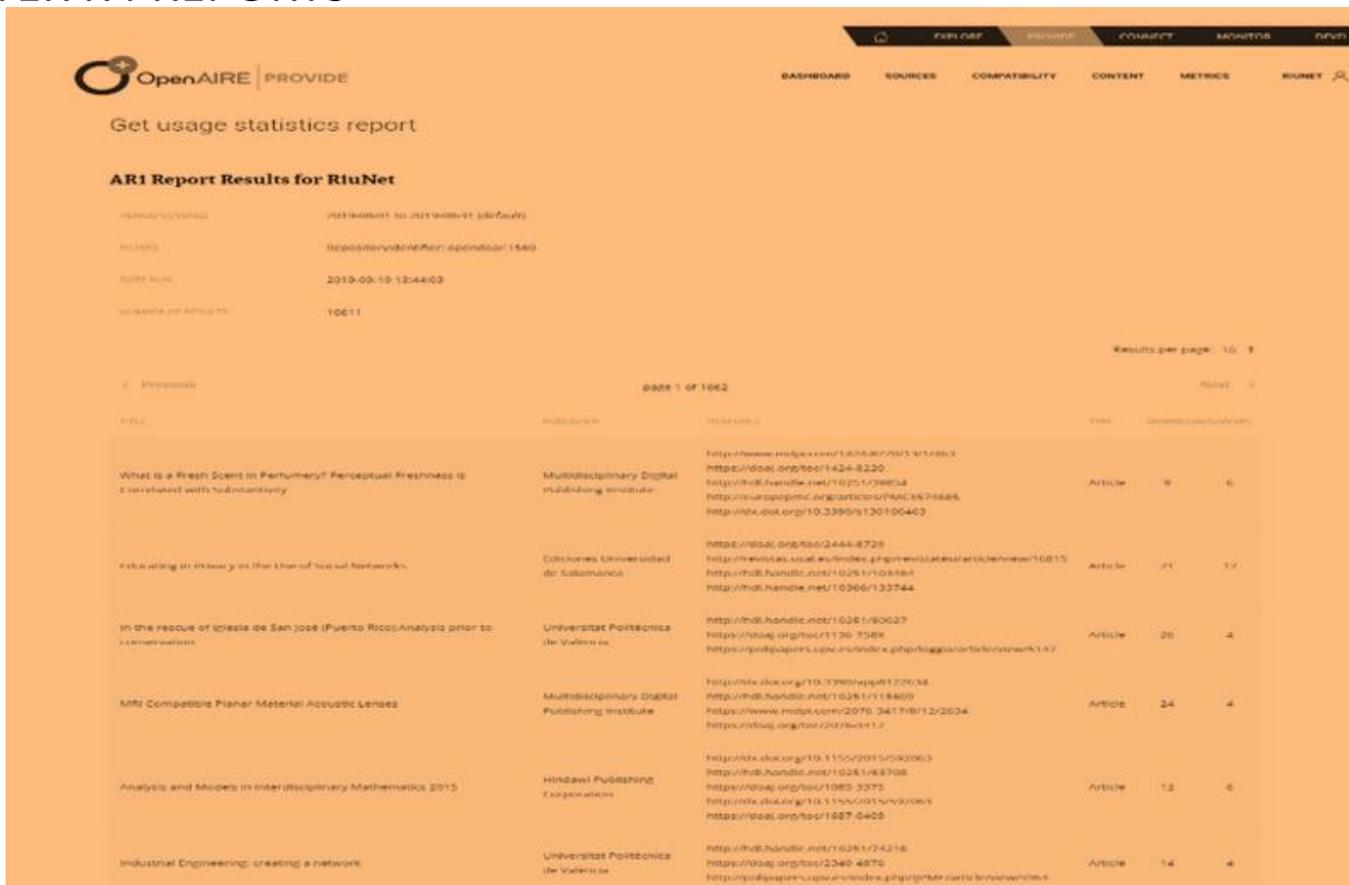
Supported Reports

 <p>AR1</p> <p>Article Report 1, number of successful article download requests by month and repository.</p>	 <p>IR1</p> <p>Item Report 1, number of successful item download requests by month and repository.</p>	 <p>RR1</p> <p>Repository Report 1, number of successful item downloads for all repositories participating in the usage statistics service.</p>
 <p>BR1</p> <p>Book Report 1, number of successful title requests by month and title.</p>	 <p>BR2</p> <p>Book Report 2, number of successful section requests by month and title.</p>	

Métricas y Estadísticas en Repositorios

Dashboard OpenAIRE

COUNTER R4 REPORTS



OpenAIRE PROVIDE

Get usage statistics report

ARI Report Results for RiuNet

REPORT ID: 20190919-10-1244-03 (default)

URL: repositoryvidentis.com/ponedor/1840

REPORT DATE: 2019-09-19 12:44:03

NUMBER OF ARTICLES: 16611

Results per page: 10

page 1 of 1662

title	publisher	url	type	articles	views
What is a Fresh Scent in Perfumery? Perceptual Freshness is correlated with substantivity	Multidisciplinary Digital Publishing Institute	http://www.mdpi.com/1422-0067/21/3/1199/1 https://doi.org/10.3390/1424-8220 http://dx.doi.org/10.3390/1424-8220 http://arxiv.org/abs/1903.04664 http://dx.doi.org/10.3390/130100463	Article	9	6
Fabrating in History in the Use of Social Networks	Ediciones Universidad de Salamanca	https://doi.org/10.1016/j.elsevier.2019.05.001 http://www.elsevier.com/locate/elsevier http://dx.doi.org/10.1016/j.elsevier.2019.05.001 http://dx.doi.org/10.1016/j.elsevier.2019.05.001	Article	21	12
In the rescue of Iglesia de San José (Puerto Rico): Analysis prior to conservation	Universitat Politècnica de València	http://dx.doi.org/10.1016/j.elsevier.2019.05.001 https://doi.org/10.1016/j.elsevier.2019.05.001 https://doi.org/10.1016/j.elsevier.2019.05.001	Article	26	4
ARFI Compatible Planar Material Acoustic Lenses	Multidisciplinary Digital Publishing Institute	https://doi.org/10.3390/130100463 http://dx.doi.org/10.3390/130100463 https://www.mdpi.com/2076-3417/12/2024 https://doi.org/10.3390/130100463	Article	24	4
Analysis and Models in Interdisciplinary Mathematics 2015	Hindawi Publishing Corporation	https://doi.org/10.1155/2015/592063 http://dx.doi.org/10.1155/2015/592063 https://doi.org/10.1155/2015/592063 https://doi.org/10.1155/2015/592063	Article	12	6
Industrial Engineering: creating a network	Universitat Politècnica de València	http://dx.doi.org/10.1016/j.elsevier.2019.05.001 https://doi.org/10.1016/j.elsevier.2019.05.001 https://doi.org/10.1016/j.elsevier.2019.05.001	Article	14	4

Métricas y Estadísticas en Repositorios

Dashboard OpenAIRE

COUNTER R5 REPORTS

- [20th OpenAIRE Content Providers Community Call](#)
- COUNTER R5 en pruebas https://beta.services.openaire.eu/usagestats_r5/

Métricas y Estadísticas en Repositorios

UsageCounts Portal OpenAIRE

SUSHI-LITE REPORTS

<http://services.openaire.eu/usagestats/sushilite>

- Devuelve un fichero JSON (formato para el intercambio de datos, similar a XML)
- Tratamiento posterior con APIs en distintos lenguajes como Javascript

Report Request

Report Name: AR1

Release: 4

Requestor: anonymous



Report Filters

DATE RANGE

Valid date formats are yyyy-mm-dd or yyyy-mm. Default range is the last available month.

Begin Date: 2020-09

End Date: 2020-12

FILTERS

Provide either a Repository Identifier or an Item Identifier
Identifier format: namespace:value
Valid namespace for Repository Identifier: openaire or opendoar.
Valid namespace for Item Identifier: openaire, doi or oid(for OAI-PMH).

Repository Identifier: opendoar:912

Item Identifier: e.g. openaire:od_...:fbgode6



Report Attributes

Valid Granularity values: Monthly or Totals

Granularity: Monthly

The Pretty attribute is just for humans playing with the API and looking at results in a browser.

Pretty print json(p) for humans

GET REPORT

Métricas y Estadísticas en Repositorios

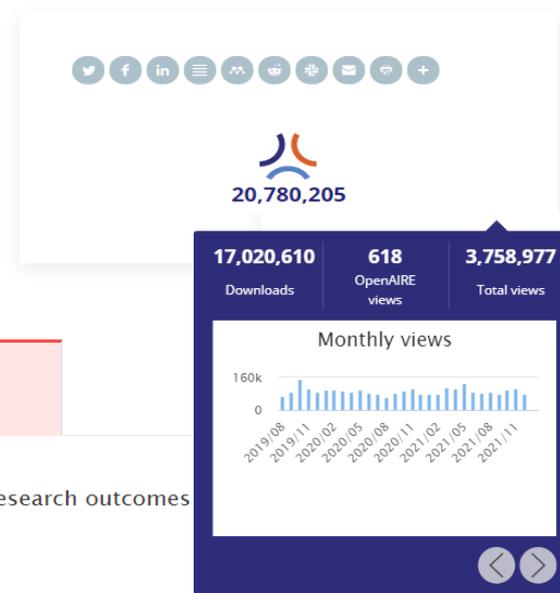
Explore OpenAIRE

Institutional Repository
e-Archivo

Repositorio Institucional de la Universidad Carlos III de Madrid

Web page: <https://e-archivo.uc3m.es/>

OPENAIRE 2.0+ (DRIVER OA, EC FUNDING) SPAIN



20,780,205

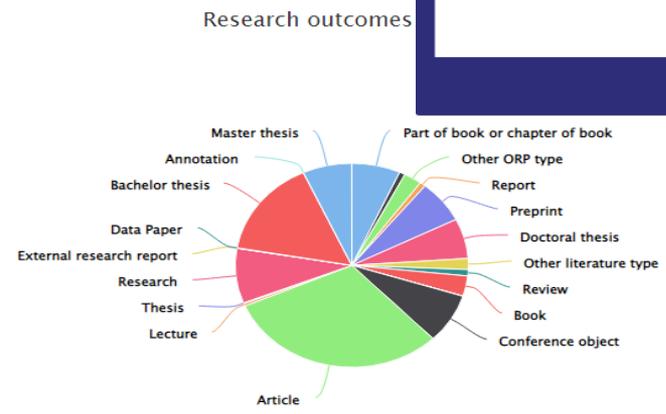
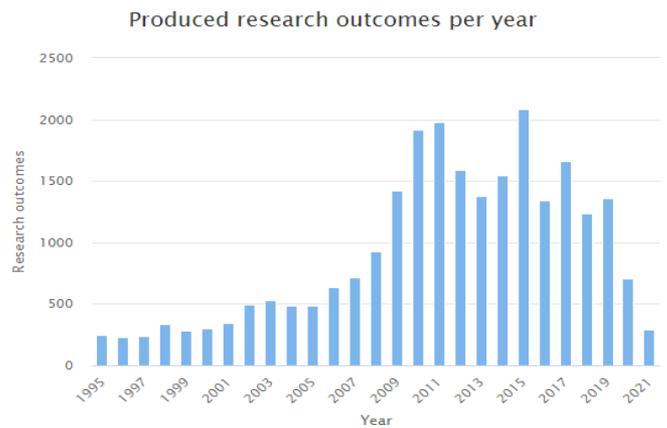
17,020,610 Downloads

618 OpenAIRE views

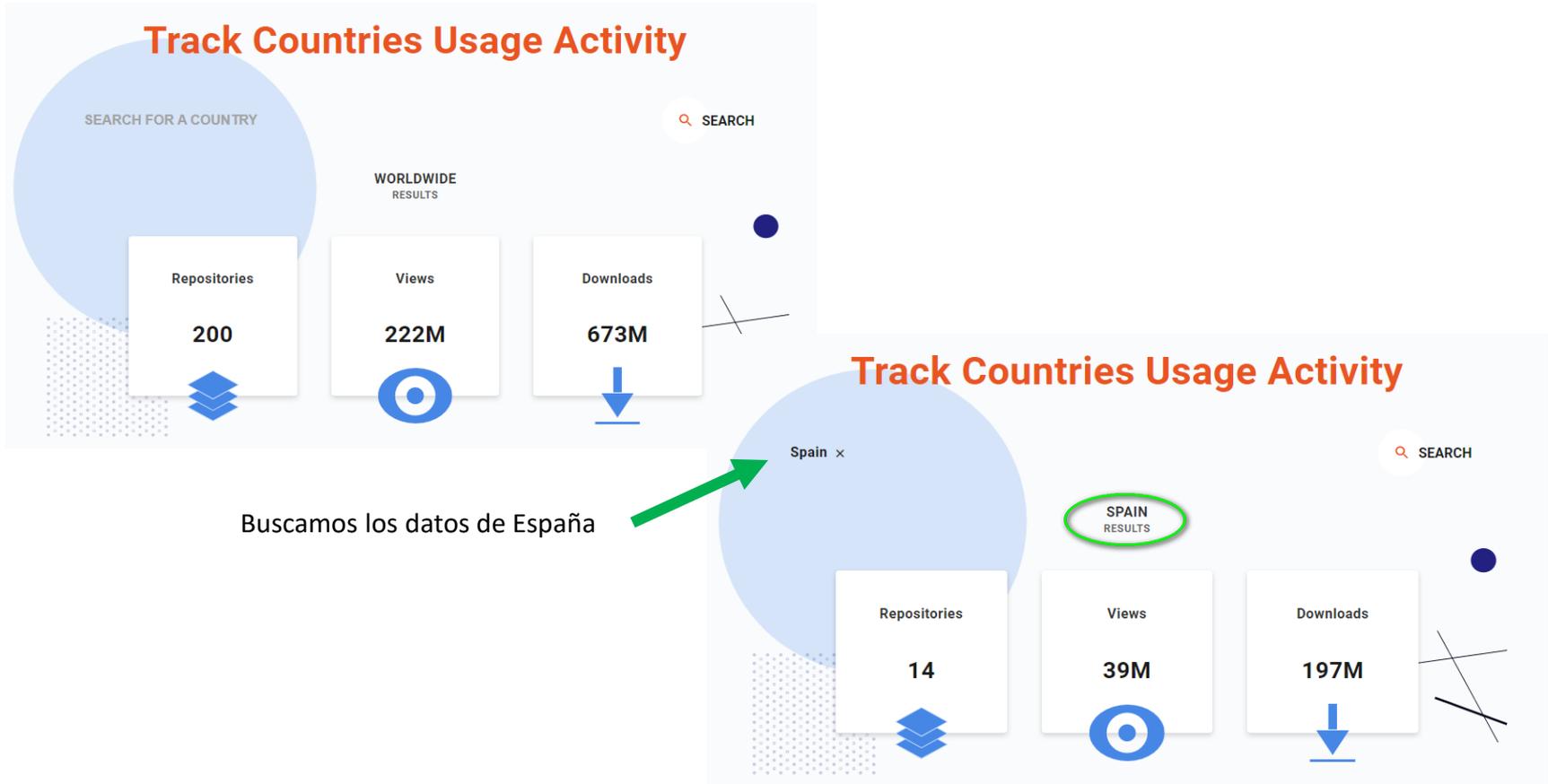
3,758,977 Total views

Monthly views

SUMMARY	RELATED CONTENT PROVIDERS	PUBLICATIONS	RESEARCH DATA	OTHER RESEARCH	STATISTICS
		25,830	1	899	



Métricas y Estadísticas en Repositorios UsageCounts OpenAIRE



Métricas y Estadísticas en Repositorios

Matomo Analytics en OpenAIRE

Repositorios españoles

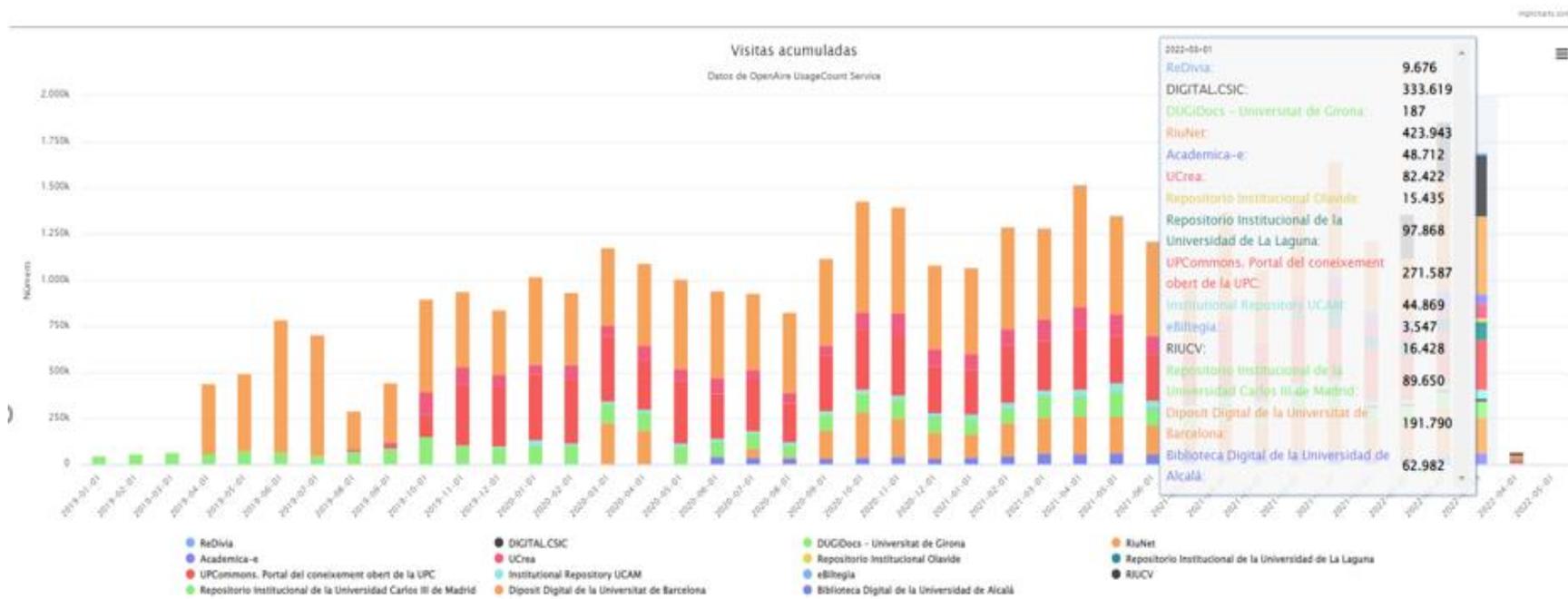
15 repositorios con el plugin de Matomo de OpenAIRE instalado y enviando datos

2018	2019	2020	2021	2022
e-Archivo (UC3M)	RiuNet (UPV)	Dispòsit Digital (UB)	ReDivia (IVIA)	DIGITAL.CSIC (CSIC)
	UCrea (UniCan)	Biblioteca Digital (UAH)	RIUCV (UCV)	Repositorio (UPO)
	UPCommons (UPC)	eBiltegia (U.Mondragón)	DUGiDocs (UdG)	
	Repositorio (UCAM)		Academia-e (UPNA)	
			Repositorio (ULL)	

Métricas y Estadísticas en Repositorios

OpenAIRE UsageCount Service

Descargas y visitas registradas de repositorios españoles

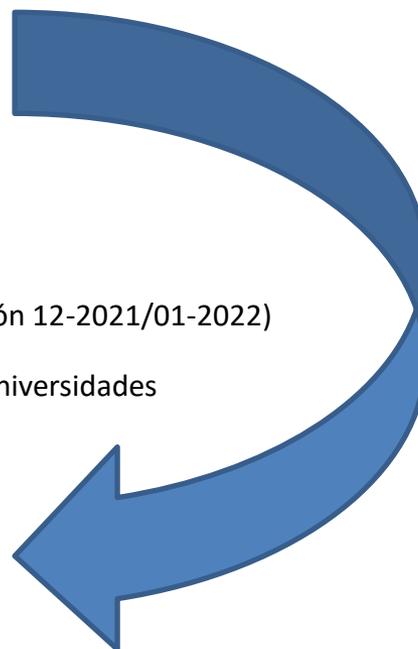


Métricas y Estadísticas en Repositorios

Grupo de Trabajo de Repositorios REBIUN

Problemas del método PUSH de Matomo

- ❖ Elevado número de eventos recogidos
- ❖ Procesamiento lento de los datos
- ❖ Desfase en la presentación de resultados (última actualización 12-2021/01-2022)
- ❖ Mantenimiento de OpenAIRE con voluntarios de distintas universidades



Solución



LA Referencia
Red de repositorios de acceso abierto a la ciencia

- ❖ Nuevo procedimiento de envío de datos a OpenAIRE con el método PULL
- ❖ Análisis de datos y presentación de resultados con SW complementario

Métricas y Estadísticas en Repositorios

Grupo de Trabajo de Repositorios REBIUN

Tareas de difusión y resultados

- ✓ Workshop REBIUN 2020: Tres experiencias con OpenAIRE: [La experiencia de la Universidad Carlos III de Madrid con métricas y estadísticas de OpenAIRE](#)
- ✓ Colaboración en la 4ª [Guía para la evaluación de Repositorios institucionales de Investigación](#)

6.- LOGS Y ESTADÍSTICAS

6.1.- Los logs del servidor web donde está alojado el repositorio se archivan de forma permanente.

Resultado:

- Sí
- No

6.2.- El repositorio proporciona un servicio de estadísticas sobre el uso de los recursos almacenados.

Se proporcionan de forma pública datos de accesos y descargas de forma individualizada para cada documento almacenado.

Resultado:

- Sí
- No

6.3.- Se realiza un filtrado de accesos de los robots o motores de búsqueda.

Resultado:

- Sí
- No

6.4.- Se realiza un filtrado de doble clics.

Resultado:

- Sí
- No

6.5. Estadísticas públicas.

El repositorio ofrece en un lugar visible estadísticas públicas anuales, al menos, sobre los siguientes aspectos: evolución de los contenidos, descargas y número de items en acceso abierto.

Resultado:

- Sí
- No

6.6. Anonimización de logs.

El repositorio es capaz de enviar los logs del servidor web de forma anonimizada para su posterior tratamiento y/o envío a recolectores de estadísticas.

Resultado:

- Sí
- No

6.7. Utilización del estándar COUNTER.

El repositorio utiliza un software de estadísticas capaz de aplicar los criterios COUNTER para una contabilización homogénea de los accesos y descargas de los registros.

Resultado:

- Sí
- No

6.8. Recolectores de estadísticas.

Se facilitan los datos a recolectores estadísticos regionales, nacionales o internacionales para una visión general de uso de los registros. Iniciativas como por ejemplo la del Consorcio Madroño, Recolecta u OpenAIRE respectivamente.

Resultado:

- Sí
- No

Métricas y Estadísticas en Repositorios

Grupo de Trabajo de Repositorios REBIUN

- ✓ [Métricas y Estadísticas en Repositorios Españoles](#). Latmetrics 2021 (Póster y vídeo explicativo)
- ✓ [Manual de Matomo Analytics y su integración con OpenAIRE para el envío de datos estadísticos](#)
- ✓ Jornadas 15 de años de RECOLECTA

Y seguimos trabajando...

Síguenos:

 <https://www.facebook.com/fecyt.ciencia>

 https://twitter.com/FECYT_Ciencia

 https://www.instagram.com/fecyt_ciencia

 <https://www.youtube.com/FECYT>

GRACIAS