



EOSC-IF

Research Products Accounting Architecture and Interoperability Guidelines

The EOSC Future project is co-funded by the European Union Horizon Programme call INFRAEOSC-03-2020, Grant Agreement number 101017536



Version 1.0
2023 March

EOSC-IF / Research Products Accounting Architecture and Interoperability Guidelines

Lead by **OpenAIRE AMKE**

Authored by Dimitris Pierrakos (OpenAIRE AMKE), Paolo Manghi (OpenAIRE AMKE)

Reviewed by QRP₁ (organisation) & QRP₂ (organisation)


Dissemination Level of the Document

Restricted to a group specified by the consortium (including the European Commission Services)

Abstract

The guidelines are aimed to provide orientation for EOSC data source managers about participation in the Research Products Accounting Service and about the methods and standards used to collect and process usage data in order to generate comparable, standards-based usage statistics. The guidelines follow the Release 4 and Release 5 of the COUNTER Code of Practice for e-Resources.

Version History

Version	Date	Authors/Contributors	Description
V0.1	24/10/2222	Dimitris Pierrakos (OpenAIRE AMKE), Paolo Manghi (OpenAIRE AMKE)	Initiation – Proposed ToC – First draft
V0.2	22/03/2023	Dimitris Perrakos (OpenAIRE AMKE) Andreas Czerniak  (Bielefeld University Library)	
V0.3			
V0.4			
V0.5			
V0.6			
V1.0		Author 1(organisation), Author 2 (organisation),Ron Dekker (TGB), Mike Chatzopoulos (ATHENA)	Final Version submitted to EC

Copyright Notice



This work by Parties of the EOSC Future Consortium is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/). The EOSC Future project is co-funded by the European Union Horizon Programme call INFRAEOSC-03-2020, Grant Agreement number 101017536.

Table of Contents

EOSC-IF / Research Products Accounting Architecture and Interoperability Guidelines	0
Glossary	4
List of Abbreviations	5
Note for authors of EOSC-Core Interoperability Guidelines:	6
1 Intended Audience	6
2 Description and main features	6
3 Response to Community Need	6
4 High-level Service Architecture	6
5 Definitions	7
6 Licensing Information	7
7 Related Guidelines	7
8 Adopted Standards	7
9 Integration Procedure	8
10 Interoperability Guidelines	8
10.1 Guidelines for Usage Tracking	8
10.1.1 PUSH Workflow	8
10.1.2 PULL Workflow	10
10.2 Guidelines for Usage Statistics Consuming via COUNTER Reports	11
11 Examples of solutions implementing this specification	13
References	14

Table of Tables

Table 1: List with references of the main standards and protocols/APIs adopted by this service

Table 2: Table 10-1: Information required to enable Usage Tracking at UsageCounts service

Table 3: Description of Master reports for filtering and customized views

Table 4: Information required to retrieve COUNTER reports from Content Providers

Table 5: Information required to retrieve COUNTER Reports from UsageCounts Service

Table of Figures

Figure 1: Service Architecture

Figure 2: Example of a report for COUNTER CoP R5

Figure 3: Zenodo (zenodo.org) since January 2023

Figure 4: B2Share (b2share.eudat.eu) since January 2023

Figure 5: Research product with usage statistics

Figure 6: Research product with usage statistics expanded

Glossary

EOSC Future project Glossary is incorporated by reference: <https://wiki.eoscfuture.eu/x/JQCK>

List of Abbreviations

Acronym	Definition
API	Application Programming Interface
COUNTER CoP	COUNTER Code of Practice
EOSC	European Open Science Cloud

1. Intended Audience

This interoperability guidelines are intended for adoption by *Open Science Data Source managers* and *Research Infrastructures administrators*, who should employ them in order to achieve an accurate and objective performance monitoring of the data source content, as well as evidenced and comparable based usage metrics (like views, downloads) of research aggregated from all over the world. *Researchers*, *Funding Organizations* and *Policy Makers*, could also exploit the guidelines to retrieve easily, fast and dynamically the research uptake, identify open science trends based on user activity, examine the engagement, and evolution or research popularity and evaluate the results of invested funds and policies.

2. Description and main features

The EOSC Research Products Accounting is operated by OpenAIRE's UsageCounts service¹. The service collects usage activity from events from EOSC Data sources, i.e. EOSC services that host collections of research products, like articles, books, datasets, etc. and include data repositories, software repositories, and publication repositories. The service forms metrics of usage activity of these Data sources, categorizing the data retrieved by number of downloads, number of views, number of repositories and all derivative quantitative open metrics, comprehensively. UsageCounts service provides standards for usage data exchange, it complies to the COUNTER Code of Practice for reliable and comparable reports, it respects user's privacy via IP anonymization of usage events, it offers global coverage and enables accumulation of usage for the same research products by exploiting the metadata deduplication functionality of the EOSC Research Graph.

3. Response to Community Need

The EOSC Research Products Accounting service covers the needs of content providers and consumers and offers added value to assist them in reaching their goals.

For data source providers, the service offers accurate and objective performance measures monitoring of the data source content, evidence based analytical metrics of views and downloads, aggregated from all over the world and activity and trends of open science topics. For the consumers the service provides trustable, accurate statistics of open science research and evaluates the result of invested funds, policies, measures and view impact.

4. High-level Service Architecture

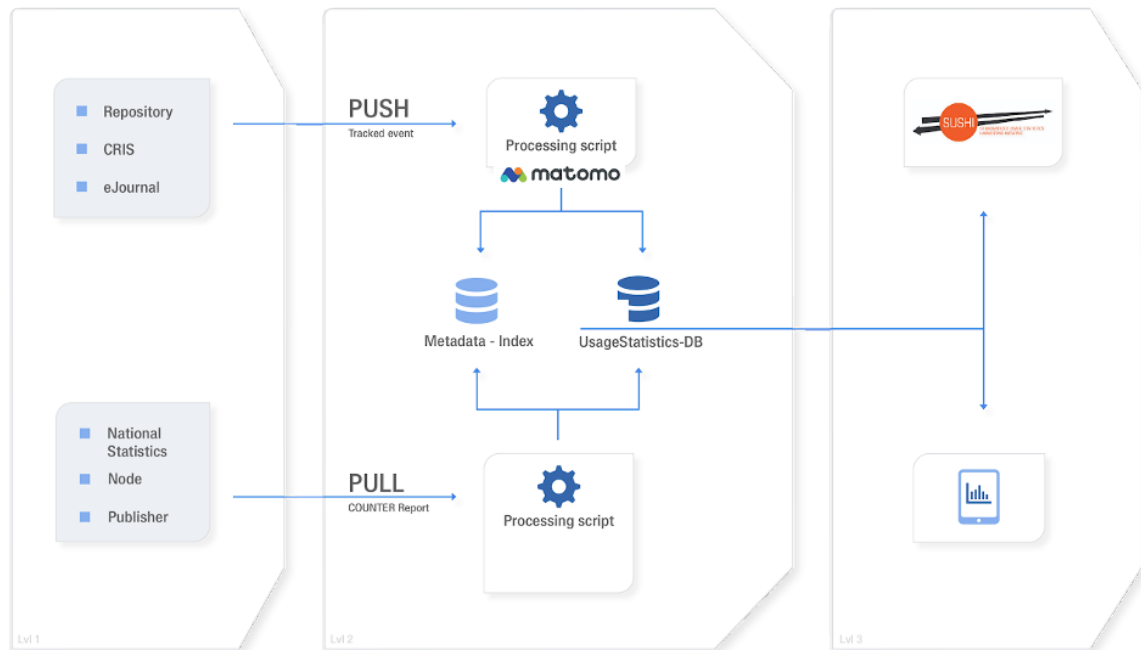
Service architecture comprises two approaches or workflows:

- A *PUSH* Workflow which allows server side real-time tracking using platform specific tracking software or using a generic log file parser based on Python that parses web server log files. Usage events are dispatched to the Matomo Analytics platform by exploiting the platform's API. The PUSH workflow supports anonymization of IPs.

¹ <https://usagecounts.openaire.eu>

- A *PULL* Workflow that collects COUNTER CoP usage statistics reports. The reports are retrieved using the Sushi-Lite API².

Figure 1: Service Architecture



5. Definitions

- *COUNTER Code of Practice*: A standard that ensures that vendors and publishers can provide their customers with consistent, credible and comparable usage data.
- *Matomo*: the most common free and open-source web analytics application to track online visits to one or more websites and display reports on these visits for analysis.
- *View*: a visit to the research product's metadata
- *Download*: the download of the research product's full text.

6. Licensing Information

CC ATTRIBUTION 4.0 INTERNATIONAL LICENSE³

7. Related Guidelines

- *OpenAIRE Guidelines*⁴
- *COUNTER CoP R4*⁵

² https://app.swaggerhub.com/apis/devhatem/COUNTER_SUSHI5_API/1.0.0

³ <https://creativecommons.org/licenses/by/4.0/>

⁴ <https://guidelines.openaire.eu/en/latest/>

⁵ <https://www.projectcounter.org/code-of-practice-sections/archived-code-of-practice-release-4/>

- COUNTER CoP R5⁶
- COUNTER CoP for Research Data⁷

8. Adopted Standards

Table 1: List with references of the main standards and protocols/APIs adopted by this service

Protocol/API	Short Description	References
Matomo Tracking API	UsageCounts uses Matomo's Tracking API to track research product's metadata views and downloads. It sends HTTP requests (GET or POST) to the service's tracking HTTP API endpoint with the correct query parameters set.	https://developer.matomo.org/api-reference/tracking-api
COUNTER_SUSHI API	The COUNTER_SUSHI API represents a RESTful implementation of SUSHI automation intended to return COUNTER reports and snippets of COUNTER usage in JSON format	https://app.swaggerhub.com/apis/COUNTER/counter-sushi_5_0_api/1.0.0
RESEARCH_DATA_SUSHI API	RESEARCH_DATA_SUSHI API represents a RESTful implementation of SUSHI automation intended to return COUNTER Research Data Release 1 reports and snippets of usage in JSON format	https://app.swaggerhub.com/apis/COUNTER/researchdata-sushi_1_0_api/1.0.0

9. Integration Procedure

Integration with the UsageCounts Service, comprises the following steps for EOSC Providers:

PUSH Workflow:

1. registration of the Provider via the EOSC provider's portal or via OpenAIRE Provide.
2. registration to the UsageCounts Service on OpenAIRE provide
3. installation of the tracking code
4. tracking of usage events from the provider
5. validation of tracking of usage events from the provider
6. retrieval of usage statistics reports presented in the EOSC provider's portal or by OpenAIRE Provide, or via a SUSHI-Lite API endpoint.

PULL Workflow:

1. registration of the Provider via the EOSC provider's portal or via OpenAIRE Provide.
2. registration to the UsageCounts Service on OpenAIRE provide
3. registration of the SUSHI-Lite API provider's endpoint
4. retrieval of usage statistics reports
5. retrieval of usage statistics reports presented in the EOSC provider's portal or by OpenAIRE Provide, or via a SUSHI-Lite API endpoint.

⁶ <https://www.projectcounter.org/code-of-practice-five-zero-two/>

⁷ <https://www.projectcounter.org/code-practice-research-data/>

1. Interoperability Guidelines

The guidelines describe the parameters that should be configured and the procedures that should be followed by the data source managers to enable the UsageCounts service in order to collect usage events and deploy standards-based usage statistics.

1.1 Guidelines for Usage Tracking

1.1.1 PUSH Workflow

In order to enable the **usage tracking PUSH** workflow, the content provider should download one of the tracking software solutions below:

- A patch for various versions of [DSpace](#).
- An [Eprints](#) plugin for version 3.
- A python script in the form of a [Generic Matomo Tracker](#) for all platforms.

A set of parameters are required for the configuration of the software above, which are summarised in Table 2 below⁸.

Table 2: Information required to enable UsageTracking at UsageCounts service

Parameter Name	Definition	Type	Required	Example
trackerURL	Endpoint of the Matomo platform	String	Mandatory	https://analytics.openaire.eu/piwik.php
idSite	the ID of the provider as registered in Matomo	Integer	Mandatory	17
tokenAuth	Matomo Authorization key used to authenticate the API request	String	Mandatory	123456
ip anonymization bytes	IP Address Anonymization Bytes	Values in {1,2,3}	Mandatory	192.168.1.2 With ip anonymization bytes=2 192.168.0.0
enabled	Flag to enable/disable tracking. Default is true	Boolean	Mandatory	True
retry	Flag whether failed reports of views should be stored in a local SQLite database to resend them later	Boolean	Mandatory	True
matomoDbLocation	Location of the SQLite DB	String	Mandatory	var/Matomo.db

⁸ <https://openaire.github.io/usage-statistics-guidelines/>

In order to enable the **usage tracking PULL** workflow and retrieve the above reports from Content Providers, the UsageCounts service has to configure a set of parameters which are summarised in the Table 4 below.

Table 4: Information required to retrieve COUNTER reports from Content Providers

Parameter Name	Definition	Type	Required	Example
baseURL	Endpoint of the COUNTER Reports	String	Mandatory	https://irus.jisc.ac.uk/api/sushilite/v1_7/
Release	Release of the COUNTER CoP	Integer	Mandatory	5
begin_date	Begin day (in month) to collect the reports	String	Mandatory	2022-03
end_date	End day (in month) to collect the reports	String	Mandatory	2022-05
ProviderIdentifier	Content Provider's Unique Identifier	String	Mandatory	opendoar:301

1.2 Guidelines for Usage Statistics Consuming via COUNTER Reports

To consume the COUNTER CoP Usage Reports from the UsageCounts service, a set of parameters should be specified, which are summarised in the Table below.

Table 5: Information required to retrieve COUNTER Reports from UsageCounts Service

Parameter Name	Definition	Type	Required	Example
baseURL	Endpoint of the COUNTER Reports	String	Mandatory	services.openaire.eu/sushilite/
Release	Release of the COUNTER CoP	Integer	Mandatory	4
begin_date	Begin day (in month) to collect the reports	String	Mandatory	2022-03
end_date	End day (in month) to collect the reports	String	Mandatory	2022-05
ProviderIdentifier	Content Provider's Unique Identifier	String	Mandatory	opendoar:301
ItemDataType	Data Type of the requested research product	String	Optional	Article
itemIdentifier	DOI or OID PID of the requested product	String	Optional	doi:10.xyz/12345
MetricType	COUNTER CoP R5 and COUNTER CoP for datasets Metric types	String	Mandatory	total_item_investigations
Granularity	Reports Granularity	String	Mandatory	Monthly

An example of a report for COUNTER CoP R5 for content Provider Universidade do Minho: RepositoriUM is given in the Table below:

Figure 2: Example of a report for COUNTER CoP R5

{

```

"Report_Header" : {
  "Created" : "2022-06-24T08:02:03Z",
  "Created_By" : "OpenAIRE Usage Counts Service",
  "Customer_ID" : "anonymous",
  "Report_ID" : "PR",
  "Report_Name" : "Platform Master Report",
  "Institution_Name" : "Universidade do Minho: RepositoriUM",
  "Institution_ID" : [ {
    "Type" : "Openaire",
    "Value" : "opendoar____:8e98d81f8217304975ccb23337bb5761"
  } ],
  "Report_Filters" : [ {
    "Name" : "BeginDate",
    "Value" : "2022-01"
  }, {
    "Name" : "EndDate",
    "Value" : "2022-02"
  } ],
  "Report_Attributes" : [ {
    "Name" : "Attributes_To_Show",
    "Value" : "Data_Type|Access_Method"
  } ]
},
"Report_Items" : [ {
  "Platform" : "Universidade do Minho: RepositoriUM",
  "Data_Type" : "Article",
  "Access_Method" : "Regular",
  "Performance" : [ {
    "Period" : {
      "Begin_Date" : "2022-01-01",
      "End_Date" : "2022-02-28"
    }
  } ],
  "Instance" : [ {
    "MetricType" : "Total_Item_Requests",

```

```

    "Count" : "23616"
  }, {
    "MetricType" : "Total_Item_Investigations",
    "Count" : "62279"
  }, {
    "MetricType" : "Unique_Item_Requests",
    "Count" : "11998"
  }, {
    "MetricType" : "Unique_Item_Investigations",
    "Count" : "20470"
  }
]
}
}
}
}

```

2. Examples of solutions implementing this specification

The OpenAIRE UsageCounts Service has been deployed successfully in two EOSC datasources and in particular Zenodo and B2SHARE. Both datasources have employed the PUSH workflow approach and they are submitting usage events to OpenAIRE's Matomo platform. In Figures 3 and 4, we present an overview of the usage events collected in these data sources.

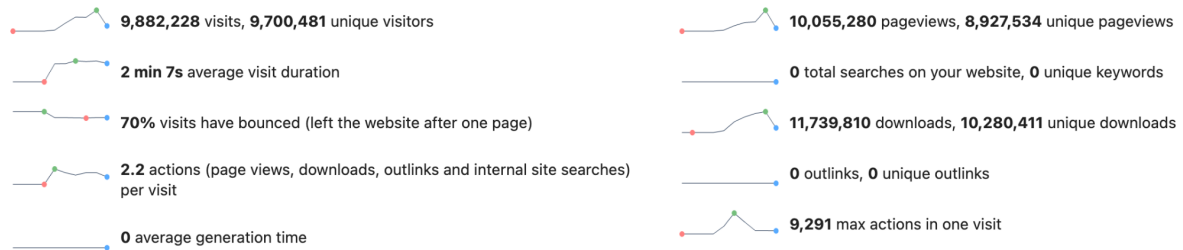
Figure 3: Zenodo (zenodo.org) since January 2022

Visits Overview



Figure 4: B2Share (b2share.eudat.eu) since January 2022

Visits Overview



In Figures 5 and 6, it is shown how the usage statistics for an individual item are published. In Figure 5, the total number is displayed, whilst in Figure 6, we show the data source or data sources from which the usage events for this particular item have been collected.

Figure 5: Research product with usage statistics





[About EOSC](#)
[Browse Marketplace](#)
[Providers Hub](#)
[Monitoring](#)

← Go to Search



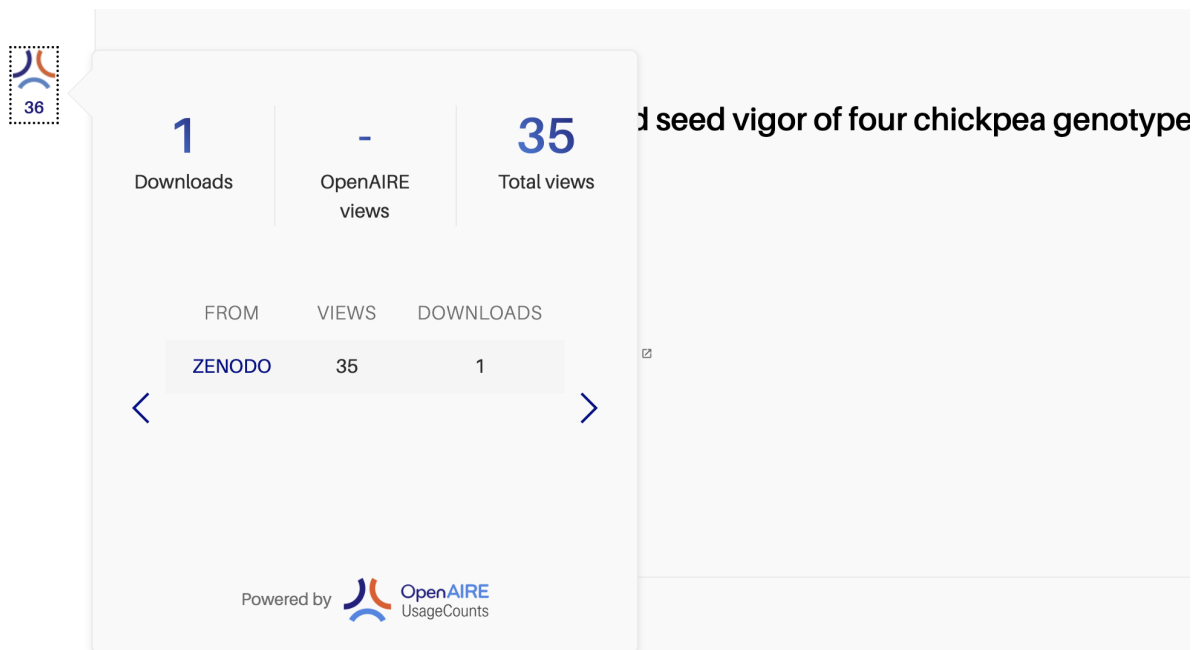
36

[View all 2 versions](#)
 Research Data . Dataset . 2021

Saline stress on germination and seed vigor of four chickpea genotypes available in Brazil

Zenodo;

Figure 6: Research product with usage statistics expanded



References

- [1] Eosc-portal.eu. 2021. EOSC Portal. [online] Available at: <<https://eosc-portal.eu/>>
- [2] Project COUNTER. <https://www.projectcounter.org/>
- [3] Dimitris Pierrakos, Andreas Czerniak, & Jochen Schirrwagen. (2020). OpenAIRE Usage Counts. The analytics service of OpenAIRE Research Graph. Research Data Alliance 16th Plenary Meeting (RDA 16th Plenary Meeting), Virtual. Zenodo. <https://doi.org/10.5281/zenodo.4268144>
- [4] Príncipe, Pedro, & Pierrakos, Dimitris. (2022, June 8). 22nd OpenAIRE Content Providers Community Call. Zenodo. <https://doi.org/10.5281/zenodo.6623488>
- [5] Schirrwagen, Jochen, Pierrakos, Dimitris, MacIntyre, Ross, Needham, Paul, Simeonov, Georgi, Príncipe, Pedro, & Dazy, André. (2017). OpenAIRE2020 - Usage Statistics Services - D8.5. Zenodo. <https://doi.org/10.5281/zenodo.1034164>