1.1 OpenAIRE Guidelines for Collecting Usage Events and Provision of Usage Statistics v1

1.1.1 Purpose

The guidelines are aimed to provide orientation for data source managers about participation in the OpenAIRE Usage Statistics 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 of the COUNTER Code of Practice for e-Resources[1] supplemented by the IRUS-UK Code of Practice[2].

1.1.2 Scope of Application

The OpenAIRE Usage Statistics Service gathers usage data and consolidated usage statistics reports respectively from its distributed network of data providers (repositories, e-journals, CRIS) by utilizing open standards and protocols and exploiting reliable, consolidated and comparable usage metrics like counts of item downloads and metadata views conformant to COUNTER.

1.1.3 Usage Data Collection, Processing and Reporting

The OpenAIRE Usage Statistics Service supports two different ways of obtaining usage information from data providers:

1) Usage events on items in data sources, such as document landing pages and full-text files, can be tracked and will be "pushed" as raw usage data to the OpenAIRE Analytics service which is based on the open source software Piwik. The cleaning process consists of two steps. At first usage data resulting from bots, spiders and web crawlers are excluded by applying a community-maintained robot list. Subsequently the data processing rules from the COUNTER Code of Practice are applied to identify successful and valid requests, sessions, and to eliminate double clicks.
   Tracked usage events contain the following paramters:

Parameter	Description
idSite	the ID of the repository
idVisit	a visitor/session ID (an 8 byte binary string)
visitIp	the IP address of the visitor
action	the action performed (view, download, outlink, etc)
url	the url of the requested item
timestamp	the date & time of the request

[1] https://www.projectcounter.org/code-of-practice-sections/general-information/
[2] http://irus.mimas.ac.uk/help/toolbox/IRUS-UK_CoP_V1.0_May_2015.pdf

OAI-PMH Identifier	the Open Access Initiative identifier of the item being viewed/downloaded
agent	the Web Browser and the operating system of the visitor
referrer	The url that is linked to the item requested

The IP address will be anonymized in Piwik by masking the last two bytes.

2) Alternatively COUNTER conformant usage statistics can be collected from data provider endpoints that support the RESTful Standardized Usage Statistics Harvesting Initiative (SUSHI) protocol[3] [4].

Consolidated Usage Statistics will be made available in OpenAIRE by the following methods:

- In the data provider dashboard for data source managers with support of configuration updates, visualization and reporting functionalities on the data provider level
- In the OpenAIRE portal by public usage statistics visualization and reporting functionalities on the data provider and individual item level
- In the OpenAIRE-API with access to usage statistics COUNTER conformant reports supporting the RESTful SUSHI protocol.

The usage statistics made available in the OpenAIRE portal and exposed via the OpenAIRE-API are released under CC0 license.

The OpenAIRE Usage Statistics Service generates usage reports conformant to COUNTER in csv/tsv and JSON format:

- IR-1 - Item Report 1, number of successful item download requests by month and repository
- JR-1 - journal Report 1, number of successful full-text article requests by month and journal
- RR-1 - Repository Report 1, number of successful item downloads for all repositories participating in the usage statistics service

Where possible the reports will also provide the metadata views.

1.1.4 Participation and Workflow

The usage statistics workflow for data providers is as follows:

- Data provider managers who wishes to participate in OpenAIRE Usage Statistics can do so in the metrics section of the data provider dashboard. The options provided for the usage data transfer mechanism are:
    1. by usage data tracking (recommended for most repositories, CRIS)
    2. by usage data reporting (recommended for national repository statistics services; publishers of eJournals)

[3] About the SUSHI-standard: http://www.niso.org/workrooms/sushi
[4] About RESTful SUSHI-Lite: http://www.niso.org/workrooms/sushi/sushi_lite/

- In case of 1) the data provider manager is provided with a unique identifier which is the websiteId in Piwik. Information about generic and platform dependent tracking plugins are provided.
- In case of 2) the data provider manager informs OpenAIRE about the SUSHI-Lite endpoint URL from where usage data reports can be queried and downloaded
- OpenAIRE usage statistics service tracks or downloads usage data, performs cleaning operations, applies COUNTER rules, associates with corresponding metadata records in the OpenAIRE index, and generates statistics
- The data provider manager can access usage statistics in the dashboard
- Usage statistics is presented along with the publication metadata in the OpenAIRE portal
- Usage statistics is exposed via the OpenAIRE SUSHI-Lite API endpoint to 3rd party services

1.1.5 Responsibilities

Before data providers can officially participate in OpenAIRE Usage Statistics the implementation and configuration of the tracker plugins and SUSHI-Lite endpoints respectively will be tested and validated between the data provider manager and the OpenAIRE support.

Data Provider Managers

They keep the configuration information regarding usage statistics in the data provider dashboard up to date.

In case of platform software updates (e.g. by migration, new releases) which may affect existing metadata record identifiers the data provider manager informs the OpenAIRE support.

OpenAIRE

The OpenAIRE Usage Statistics service complies with the EU data protection directive[5] which will be replaced by the General Data Protection Regulation in 2018[6] with regard to the protection of personally identifiable information.

The Usage Statistics Service will generate visualizations and reports on a monthly basis. In order to conform with the COUNTER Code of Practice the data processing and cleaning rules will be maintained to comply with the lastest release of COUNTER.

1.1.6 Software Support

Information about tracker and SUSHI plugins are provided at: https://github.com/openaire/usage-statistics .

Initially tracker plugins for DSpace and EPrints repositories are available as well as an SUSHI-Lite plugin for OJS.

[5] officially Directive 95/46/EC: http://eur-lex.europa.eu/legal-content/EN/LSU/?uri=celex:31995L0046
[6] http://ec.europa.eu/justice/data-protection/

The support of further platforms will be extended.

1.1.7 Maintenance of the Guidelines

The guidelines are continuously reviewed for their validity and will be updated with regards to new releases of standards for the recording and exchange of usage statistics and new releases of OpenAIRE services that record, process, represent or expose usage statistics. Participating data provider managers will be informed accordingly.