MONITOR Services: Final Report on VA Activities

JUNE 2023

Scholarly Communication Services for EOSC users

D5. 3 – MONITOR Services: Final Report on VA Activities

Version 1.0 – Final

PUBLIC

Report on the MONITOR installations operation activity after the last reporting period, including a detailed set of indicators.

H2020-INFRAEOSC-2020-2
Grant Agreement 101017452

Document Description

D5.3 – MONITOR Services: Final Report on VA Activities	
WP5- Virtual Access to MONITOR portfolio	
WP participating organizations: UNIWARSAW, CNR, UNIBI, ARC, UNIBO	
Contractual Delivery Date: 06/2023	Actual Delivery Date: 06/2023
Nature: Report	Version: 1.0 Final
Public Deliverable	

Preparation Slip

	Name	Organisation	Date
From	Marek Horst	UNIWARSAW	15/06/2023
Edited by	Silvio Peroni Marek Horst	UNIBO UNIWARSAW	21/06/2023 22/06/2023
Reviewed by	Alessia Bardi Alex Ioannidis Andreas Czerniak	CNR CERN UNIBI	22/06/2023 23/06/2023 26/06/2023
Approved by	Paolo Manghi	OPENAIRE AMKE	28/06/2023
For delivery	Mike Hatzopoulos	OPENAIRE AMKE	30/06/2023

Revision History

Issue	Item	Reason for Change	Author	Organization
V0.1	Draft version	The first version of deliverable.	Marek Horst	UNIWARSAW
V0.2	Draft version	Accepting an internal review outcome.	Marek Horst	UNIWARSAW
V1.0	Final version	Minor updates related to dates and versions.	Marek Horst	UNIWARSAW

Contents

Disclaimer

This document contains description of the OpenAIRE-Nexus project findings, work and products. Certain parts of it might be under partner Intellectual Property Right (IPR) rules so, prior to using its content please contact the consortium head for approval.

In case you believe that this document harms in any way IPR held by you as a person or as a representative of an entity, please do notify us immediately.

The authors of this document have taken any available measure in order for its content to be accurate, consistent and lawful. However, neither the project consortium as a whole nor the individual partners that implicitly or explicitly participated in the creation and publication of this document hold any sort of responsibility that might occur as a result of using its content.

This publication has been produced with the assistance of the European Union. The content of this publication is the sole responsibility of the OpenAIRE-Nexus consortium and can in no way be taken to reflect the views of the European Union.

OpenAIRE-Nexus is a project funded by the European Union (Grant Agreement No 101017452).

Acronyms

AAI	**Authentication and Authorisation Infrastructure**
APC	**Article Processing Charge**
BPC	**Book/Monograph Processing Charge**
DOA	**Description of Action**
GA	**Grant Agreement**
KPI	**Key Performance Indicator**
OSO	**Open Science Observatory**
RI	**Research Infrastructure**
TA	**Transformative Agreements**
UoA	**Unit of Access**
VA	**Virtual Access**

Publishable Summary

This document is the final report on the MONITOR installations operation activity after 30 months of operation, including a detailed set of indicators.

The following services were considered within the report:

- MONITOR dashboard
- OpenCitations
- ScholeXplorer
- Usage Counts
- OpenAPC
- Open Science Observatory
- OpenAIRE AAI (OpenAIRE Login)

The report is split into several sections covering services description, definition of KPIs, acquired KPI values, rationale for choosing Virtual Access indicators, acquisition methodology and measurement time period coverage. Any deviations regarding the KPIs definitions, their coverage and potential problems with reaching the success target were described in detail in the “Deviations of VA tables and KPIs” section. The "VA assessment report" section summarises an outcome of the work conducted by the *External Advisory Board for the VA assessment*.

1. Service Portfolio Description

The MONITOR portfolio includes services to monitor evolution, provenance, costs and impact of research and open science policies.

2. Definition of VA Activities

As written in the DOA, WP5 operates Virtual Access (VA) to all service installations in the MONITOR portfolio, to ensure:

- onboarding to OpenAIRE service catalogue[1] and EOSC service catalogue[2],
- quality of service,
- support to EOSC users, and
- collection of usage indicators and user accounting required to extrapolate the total quantity of access to the installations provided to users beyond their usual communities.

Virtual Access will be provided to the following installations and related reimbursement schemes:

- **MONITOR Dashboard**, deployed at ICM, unit of access "monitoring gateways", reimbursement by Unit Cost.
- **OpenCitations**, deployed at UniBo, unit of access "API requests", reimbursement by Actual Cost.
- **ScholeXplorer**, deployed at ICM, unit of access "API requests", reimbursement by Actual Cost.
- **UsageCounts**, deployed at ICM, unit of access "API requests", reimbursement by Actual Cost.
- **OpenAPC**, deployed at UniBi, unit of access "data provider institutions", reimbursement by Unit Cost.
- **Open Science Observatory**, deployed at ICM, unit of access "API requests", originally reimbursed by Actual Cost but offered in-kind.
- **OpenAIRE AAI (OpenAIRE Login)**, deployed at ICM, unit of access is "number of user logins" and the "API requests" to the AA endpoints, reimbursement by Actual Cost.

[1] http://catalogue.openaire.eu/search;quantity=10
[2] https://marketplace.eosc-portal.eu/

3. DESCRIPTION OF SERVICES

The MONITOR portfolio includes the following services:

Service name: MONITOR Dashboard

Description: customised portals as-a-service provided to funders, institutions and research infrastructures/initiatives (RIs) detailing research throughput, output, collaboration, impact and open science uptake. Based on the OpenAIRE Research Graph, a semi-automatically curated catalogue to serve the EOSC and the global research community, which includes all types of scholarly records (publications, data, software, other research artifacts), interlinked via citations, and maintaining provenance information about all actors involved in the research process (researchers, organisations, funders, service providers).

Website: https://monitor.openaire.eu/

Service name: OpenCitations

Description: OpenCitations infrastructure allows researchers, publishers, and research assessment agencies, to access, leverage, and monitor the evolution of scholarly link networks. It is realized by providing OpenAPI to access open citation data shared under CC0 licence.

Website: https://opencitations.net/

Service name: ScholeXplorer

Description: ScholeXplorer populates and provides access to a graph of links between dataset and literature objects, and dataset and dataset objects. It aggregates linked metadata harvested from the data sources and builds a harmonised and de-duplicated graph of scholarly objects. The graph is openly accessible under CC0 licence via search REST APIs returning links in Scholix format.

Website: http://scholexplorer.openaire.eu/

Service name: UsageCounts

Description: an open analytics service aggregating and de-duplicating publications DOI usage data. UsageCounts forms metrics of usage activity of Open Access Repositories categorising the data retrieved by country, number of downloads, number of views, number of repositories and all derivative quantitative open metrics, comprehensively.

Website: https://usagecounts.openaire.eu/

Service name: OpenAPC

Description: The OpenAPC initiative collects and disseminates datasets on fees paid for open access, publishing on GitHub under an open database licence. OpenAPC aggregates data on Open

Access (OA) journal articles (APCs), Open Access Books (BPCs) and data on articles published under transformative agreements (TAs), like Springer Compact or Wiley DEAL. All data is provided voluntarily by universities and other HEI, funders or national consortia.

Website: https://openapc.net/

Service name: Open Science Observatory

Description: Open Science Observatory is a dashboard that includes statistics and monitoring information on open science (openness – FAIRness - collaboration) in Europe, with country views (offered in-kind). It presents a collection of indicators and visualisations that help interested stakeholders (policy makers and research administrators among others) better understand the Open Science landscape in Europe across countries. The platform assists the monitoring, and consequently the enhancing, of open science policy uptake across different dimensions of interest, revealing weak spots and hidden potential.

Website: https://osobservatory.openaire.eu/

Service name: OpenAIRE AAI (OpenAIRE Login)

Description: OpenAIRE Login enables researchers to securely access and share common resources and services using identities from eduGAIN, the global network of academic identity federations. For users coming from the industry or citizen scientists who may not have access to eduGAIN, the OpenAIRE Login service supports additional trusted authentication providers, such as social networks, community identity providers, and other platforms such as ORCID that can provide federated user identities. OpenAIRE Login allows connecting services using popular protocols, such as OpenID Connect and SAML, to securely authenticate and identify their users, organise them in groups, assign them roles, and centrally manage access rights for accessing protected resources.

4. DEFINITION OF KPIS

Service name: OpenAIRE MONITOR

KPI name	Definition
#MonitoringGateways (Unit of Access)	Number of monitoring gateways for funders, institutions and research initiatives. It captures the effort required to ensure the specific portion of the Graph related to the requesting EOSC user can be identified, kept up-to-date, and finally delivered to Gateway users.
#IndexedRecordsInTheOpenAIREGraph	Number of records available in the OpenAIRE Graph covering publications, datasets, software and other research products
#Uptime	Service availability, expressed as percentage value

Service name: OpenCitations

KPI name	Definition
#APIrequests (Unit of Access)	Number of requests handled by publicly available APIs
#IndexedRecords	Indexed citation records
#HarvestedDataSources	Harvested data sources
#Uptime	Service availability, expressed as percentage value

Service name: ScholeXplorer

KPI name	Definition
#APIrequests (Unit of Access)	Number of requests handled by publicly available APIs
#IndexedRecords	Indexed dataset-dataset and dataset-publication relation records
#Uptime	Service availability, expressed as percentage value

Service name: UsageCounts

KPI name	Definition
#APIrequests (Unit of Access)	Number of requests handled by publicly available APIs
#RepositoriesRegistered	Repositories registered in UsageCounts
#Uptime	Service availability, expressed as percentage value

Service name: OpenAPC

KPI name	Definition
#DataProviderInstitutions (Unit of Access)	Number of institutions providing the data
#APIrequests	Number of requests handled by publicly available APIs
#Uptime	Service availability, expressed as percentage value

Service name: Open Science Observatory

KPI name	Definition
#WebVisits (Unit of Access)	Registered web visits on OSO site
#Uptime	Service availability, expressed as percentage value

Service name: AAI (Login)

KPI name	Definition
#UserLogins (Unit of Access)	User login operations issued against AAI API
#APIrequests (Unit of Access)	Number of requests handled by publicly available APIs
#RegisteredUsers	Number of registered users
#Uptime	Service availability, expressed as percentage value

5. KPIs PER SERVICE

Acquired values for indicators

This section presents all the indicator values acquired at M12, M24 and M30. The detailed description on how indicators were gathered is available in the "Indicators acquisition methodology" section. The description of the table columns is available below the KPIs table.

Service	KPI	Point of reference (Basis)	Success Target (% inc)	Initial value	Success Target value (M30)	M12 value	M24 value	M30 value
MONITOR	**#MonitoringGateways (UoA)**	25	72%	0	**18**	6	7	**20**
MONITOR	**#IndexedRecordsIn TheOpenAIREGraph**	35,939,015	30%	146,618,533	**157,400,238**	153,602,997	174,319,336	**232,166,635**
MONITOR	**#Uptime**				**99,98%**	100%	100%	**100%**
OpenCitations	**#APIrequests (UoA)**	6,500,000	138%	0	**8,970,000**	22,123,939	54,119,079	**87,145,632**
OpenCitations	**#IndexedRecords**	445,826,118	30%	759,516,507	**893,264,342**	1,235,170,583	2,647,672,913	**3,212,028,433**
OpenCitations	**#HarvestedDataSources**	4	100%	4	**8**	4	7	**8**
OpenCitations	**#Uptime**				**99,98%**	100%	100%	**99,89%**
ScholeXplorer	**#APIrequests (UoA)**	555,000,000	160%	0	**888,000,000**	510,300,371	1,108,075,737	**1,329,926,820**
ScholeXplorer	**#IndexedRecords**	260,000,000	30%	387,577,328	**465,577,328**	408,141,644	337,450,956	**337,966,658**[3]
ScholeXplorer	**#Uptime**				**99,98%**	100%	100%	~~**98.52%**~~ **100%**[4]
UsageCounts	**#APIrequests (UoA)**	649,779	115%	0	**747,246**	495,800	5,598,233[5]	**8,945,126**
UsageCounts	**#RepositoriesRegistered**	200	250%	0	**500**	436	506	**563**
UsageCounts	**#Uptime**				**99,98%**	100%	100%	**100%**

[3] The significant drop in the number of indexed records is caused by the decreased number of relationships provided by the main Datacite endpoint providing the metadata records. Relations were moved to the EventData endpoint, which is going to be covered by the OpenAIRE import procedure in the upcoming data provision round.

[4] ScholeXplorer API availability was 100% in the reported period even though checkme.openaire.eu availability monitoring tool reported 98.52%. API was fully functional at the time of alleged outage (starting from 2023-03-21 until 2023-04-03) and the reason for the false downtime indication was improper swagger-focused endpoint being monitored instead of a dedicated, newly developed health-check endpoint. An internal monitoring, which was set up for potential service restarting purposes in case of malfunction, confirmed full availability.

[5] The substantial increase of API hits is mostly bound to the methodology change described in the "Indicators acquisition methodology" section below. The major factor is including the UsageCounts HTTP GET requests issued against StatsTool API. This change was reflected also in the calculation of the "point of reference (basis)".

OpenAPC	**#DataProviderInstitutions (UoA)**	230	59%	0	**136**	93	152	**163**[6]
OpenAPC	**#APIrequests**	70,000	50%	0	**35,000**	33,795	148,016	**261,008**
OpenAPC	**#Uptime**				**99,98%**	100%	100%	**99,99%**
Observatory	**#WebVisits (UoA)**	0	in-kind	0	**0**	1339	3806	**5066**
Observatory	**#Uptime**				**99,98%**	100%	100%	**100%**
AAI	**#UserLogins (UoA)**	20,000	100%	0	**20,000**	23,260	57,855	**82,339**
AAI	**#APIrequests (UoA)**	60,000	100%	0	**60,000**	518,693	1,076,839	**1,330,010**
AAI	**#RegisteredUsers**	4,500	100%	0	**4,500**	18,407	28,161	**33,504**
AAI	**#Uptime**				**99,98%**	100%	100%	**100%**

Column description:

Service	Service name
KPI	Key Performance Indicator, marked with "UoA" if Unit of Access
Point of reference (Basis)	Base value as of 2019 provided in the Grant Agreement document, point of reference when applying percentage increase as success target
#Uptime	Service availability, expressed as percentage value
Success Target (% increase)	Success target, expressed as percentage and declared in the Grant Agreement document, to be applied on the point of reference in order to calculate success target value at M30
Initial value	Initial tracked value obtained on 2021-01-01 when the OpenAIRE-Nexus project started. Required to calculate success target value.
Success Target value (M30)	Success target value at M30. The value is either explicitly defined in the Grant Agreement document or calculated with the following formula: Initial value + (point of reference * success target)
M12 value	KPI value readout at M12 (2021-12-31)
M24 value	KPI value readout at M24 (2022-12-31)
M30 value	KPI value readout at M30 (usually 2023-06-16, numbers were gathered before the deliverable submission deadline set to 2023-06-30)

[6] Correction due to the changes in the baseline values, described in the "Indicators acquisition methodology" section below.

The rationale for choosing VA indicators

This section provides explanations on the VA KPIs of each service of the MONITOR portfolio.

Service name: MONITOR Dashboard

The unit of access for OpenAIRE MONITOR is the "Monitoring Gateway", which proportionally captures the effort required to ensure the specific portion of the Graph related to the requesting EOSC user (funder, institution, research initiative) can be identified, kept up-to-date, and finally delivered.

Service name: OpenCitations

The unit of access for the installation is the "API request", which measures an actual interest in the service and can be used to quantify the cost of service maintenance and operation. Virtual Access is provided via Actual Cost as the service offers Open APIs.

Service name: ScholeXplorer

The unit of access for the installation is the "API request", which measures the effective interest in the service and can be used to quantify the cost of service maintenance and operation. Virtual Access is provided via Actual Cost as the service offers Open APIs.

Service name: UsageCounts

The unit of access for the installation is the "API request", which measures the effective interest in the service and can be used to quantify the cost of service maintenance and operation. Virtual Access is provided via Actual Cost as the service offers an open API (SUSHI-Lite).

Service name: OpenAPC

The unit of access for the installation is the "data delivering institution", e.g. a library, a university, or a funding organisation, which self-reports or provides cost information on Open Access publishing on a *regular basis*.

Service name: Open Science Observatory

The newly elected unit of access for the Open Science Observatory installation is the #WebVisits, which measures the effective interest in the service and can be used to quantify the cost of service maintenance and operation. It replaces "API request" which was originally defined in the Grant Agreement table which could not be measured due to the lack of publicly available API.

Service name: OpenAIRE AAI (OpenAIRE Login)

OpenAIRE AAI's unit of access is the "number of user logins" and the "API requests" to the AA endpoints. Due to the enabling function of the service, for all OpenAIRE-Nexus services, hence the complexity of "splitting the cost" across services, the service is under Virtual Access Actual Cost.

Indicators acquisition methodology

This section covers detailed acquisition methodology of all Virtual Access KPIs.

The OpenAIRE infrastructure offers a dedicated service to keep track of customised metrics over time. The metrics service is based on Prometheus (https://prometheus.io/) and Grafana (https://grafana.com/). Most of the OpenAIRE services in the MONITOR portfolio exploit the metrics service so that the collection of KPI values is an automated process. The type of acquisition process (manual vs automatic with Prometheus) is specified in the "other remarks" paragraph of each service.

OpenAIRE MONITOR

KPI: #MonitoringGateways (Unit of Access)

acquisition methodology: covers EC-funded funder, institutional and research initiative gateways which were deployed after 2021.01.01 (OpenAIRE-Nexus project start) and which are not overlapping with the gateways available before 2021.01.01 (speaking other words: excluding new incarnations of already available gateways). Since M15 the eligibility rules were slightly refined by covering officially requested dashboards only. This additional condition was imposed in order to have a strong proof of demand for a dashboard creation.

other remarks: Acquisition procedure is not automated, the list is approved by the product owners.

KPI: #IndexedRecordsInTheOpenAIREGraph

acquisition methodology: It is related to the total number of research records (covering publications, datasets, software, projects and other research products) provided by OpenAIRE-Explore. This MONITOR KPI is in fact provided by the OpenAIRE-Explore service which is part of the DISCOVER bundle. It is bound to the specific OpenAIRE Graph materialisation coming in the form of the SOLR index content which is backing up the publicly available OpenAIRE endpoints.

other remarks: the base value for 2019 was not defined in the GA document so it was taken from the execution logs of the last provisioning round executed in 2019. This value (36M) is significantly lower than the initial value recorded at the beginning of the project (147M) because throughout the year 2020 OpenAIRE Research Graph was extended with a large number of records coming from CrossRef, Microsoft Academic Graph and Unpaywall.

OpenCitations

KPI: #APIrequests (Unit of Access)

acquisition methodology: The total number of requests handled each month is calculated on OpenCitations side and exposed via the Prometheus exporter endpoint. Apart from the total

number of handled requests, there are metrics related to particular OpenCitations API endpoints tracked by the metrics acquisition system.

KPI: #IndexedRecords

acquisition methodology: This service deals with citations, so the number of indexed records is related to citations only.

KPI: #HarvestedDataSources

acquisition methodology: This indicator is a cumulative value and refers to the number of harvested data sources so far.

OpenAIRE ScholeXplorer

KPI: #APIrequests (Unit of Access)

acquisition methodology: Number of API requests is retrieved from the log files generated by ScholeXplorer service.

other remarks: wrongly calculated base value (for 2019) defined in the Grant Agreement document, more details in *6.Deviations of VA tables and KPIs section*.

KPI: #IndexedRecords

acquisition methodology: ScholeXplorer is all about the relations and provides the graph of links between dataset and literature objects and dataset and dataset objects. The total number of relations is exposed as the *scholexplorer_indexed_relations* metric by Prometheus exporter service, which is reported as the total number of indexed records.

other remarks: 2021 initial value actually refers to August 2021. We cannot go further back, mainly because there was significant ScholeXplorer refactoring that substantially increased the quality and decreased the number of relations. The value reported back in August 2021 is the first value which is comparable with current readouts.

OpenAIRE UsageCounts

KPI: #APIrequests (Unit of Access)

acquisition methodology: UsageCounts related requests issued against UsageCounts and StatsTool APIs are reported in the internal log files. An automated procedure was established to import those entries incrementally and fetch in batches to the OpenAIRE web analytics system available at https://analytics.openaire.eu. As a second step, another generic procedure is responsible for exposing API request metrics to Prometheus, the central point of the Virtual Access Metrics Acquisition Subsystem.

other remarks: 2019 base value was not specified in the Grant Agreement document. The Grant Agreement document mentions events related to 60M views and 200M downloads, but those numbers are related to external repository related events (tracked by usagestats) and do not reflect an uptake of UsageCounts service. The closest we could get was the one-year period,

2020.07.18 - 2021.07.17, the earliest date range available in Matomo (the OpenAIRE analytics platform).

In the beginning of the OpenAIRE-Nexus project the #APIRequests KPI covered the requests issued against UsageCounts API only because it was the only endpoint handling UsageCounts related requests.

On 11th of October 2021 Stats Tool API, which is a part of OpenAIRE API portfolio, took over the responsibility of handling the vast majority of UsageCount requests related to published usage statistics due to performance reasons. As a result of this fact Stats Tool API was covered with Virtual Access metrics monitoring and the number of UsageCount related HTTP POST requests handled by the Stats Tool API was included in the grand total number of UsageCounts related requests handled in the reporting period.

After 18th of April, 2022 the evolution of Stats Tool API continued which resulted in Stats Tool API HTTP POST requests going down from ~600 to ~50 per day while the increase was observed in Stats Tool API GET requests. Originally StatsTool GET requests were not included in the UsageCounts VA related counts because POST requests reflected the complete volume of UsageCounts API hits which were taken over by the StatsTool API on 11th of October 2021. The evolution of request types handled by Stats Tool API resulted not only in including requests of GET-type in VA related counts but also required base value recalculation because this type of requests was not included when calculating base value for the first VA report (D5.1). It did not affect base value in a significant way because formerly defined value 646,268 was replaced with 649,779 and the success target was recalculated respectively from 743,208 to 747,246.

KPI: #RepositoriesRegistered

acquisition methodology: Refers to the number of repositories that have enabled their metrics and have an assigned Matomo[7] tracking identifier. Provided by OpenAIRE-PROVIDE service.

OpenAPC

KPI: #DataProviderInstitutions (Unit of Access)

acquisition methodology: Current number of contributing institutions registered in OpenAPC since the beginning of 2021.

other remarks: In order to determine the initial values for participating institutions/organizations, UNIBI assessed the number of contributing institutions and used that information to improve the baseline count for key performance indicators (KPIs). This improvement involved considering all organizations that contributed, their integration date, and Git-MergeId from the official OpenAPC - GitHub repository[8] and took place in March 2023. As a result, the APC baseline count increased from 260 to 272, while the BPC baseline count rose from

[7] https://matomo.org/
[8] https://github.com/OpenAPC/openapc-de/

0 to 5. It is important to note that these changes have no impact on the successful VA units delivery outcome for M30.

KPI: #APIrequests

acquisition methodology: The number of requests handled each month is calculated by OpenAPC and published to the GitLab repository via the metrics exporter endpoint.

other remarks: API requests readouts are not available for the first months of the project (January 2021 - August 2021), before introducing automated metrics acquisition, due to UniBi GDPR policy. The 2019 base value was not defined in the GA document, but it was provided by UniBi in the OpenAIRE ticketing system.

Open Science Observatory

KPI: #APIrequests ~~(Unit of Access)~~

acquisition methodology: Indicator cannot be provided because there is no publicly available API which could be covered with measurement.

other remarks: OSO does not expose any API to handle requests, therefore falling back to #WebVisits as Unit of Access, more details in *6.Deviations of VA tables and KPIs section*.

KPI: #WebVisits (Unit of Access)

acquisition methodology: OSO website was integrated with Matomo tracking tool so the web visits metric is translated from unique page views provided by analytics.openaire.eu and stored in the VA metrics acquisition system.

other remarks: Unit of Access KPI was redefined to #WebVisits. There were no visits registered in 2019 because OSO was deployed in production in 2020. OSO got integrated with the OpenAIRE analytics platform (Matomo) in June 2021 so the metrics for the first half of 2021 could not be calculated.

OpenAIRE AAI (OpenAIRE Login)

KPI: #RegisteredUsers

acquisition methodology: Total number of registered users is tracked by AAI and exposed via the Prometheus exporter endpoint. This number represents state, so an increase needs to be measured in a reference to a non-zero initial value (number of registered users before 2021.01.01).

other remarks: wrongly defined as Unit of Access in the UoA table in the Grant Agreement document, more details in *6.Deviations of VA tables and KPIs section*.

KPI: #UserLogins (Unit of Access)

acquisition methodology: Total number of login operations performed by users is tracked by AAI and exposed via the Prometheus exporter endpoint. An endpoint was integrated with metrics

acquisition subsystem in May 2021 and the indicator does not cover events issued before 18th of March 2021.

KPI: #APIrequests (Unit of Access)

acquisition methodology: Total number of API requests is tracked by AAI and exposed via the Prometheus exporter endpoint. An endpoint was integrated with the metrics acquisition subsystem in May 2021 but the indicator covers events issued after 2021-01-01 (since the beginning of the OpenAIRE-Nexus project).

Time period coverage

Section below describes the metrics acquisition time period. Some indicators were not tracked since the beginning of the OpenAIRE-Nexus project because the technical solutions for an automated acquisition of those indicators were introduced after the project started.

Service name: OpenAIRE MONITOR

KPI name	From	To
#MonitoringGateways (Unit of Access)	01-01-2021	23-06-2023
#IndexedRecordsInTheOpenAIREGraph	01-01-2021	16-06-2023
#Uptime	01-01-2021	31-05-2023

Service name: OpenCitations

KPI name	From	To
#APIrequests (Unit of Access)	01-01-2021	16-06-2023
#IndexedRecords	01-01-2021	16-06-2023
#HarvestedDataSources	01-01-2021	16-06-2023
#Uptime	12-07-2021	31-05-2023

Service name: ScholeXplorer

KPI name	From	To
#APIrequests (Unit of Access)	01-01-2021	31-05-2023
#IndexedRecords	01-01-2021	16-06-2023
#Uptime	01-01-2021	31-05-2023

Service name: UsageCounts

KPI name	From	To
#APIrequests (Unit of Access)	01-01-2021	16-06-2023
#RepositoriesRegistered	01-01-2021	16-06-2023
#Uptime	01-01-2021	31-05-2023

Service name: OpenAPC

KPI name	From	To
#DataProviderInstitutions (Unit of Access)	01-01-2021	16-06-2023
#APIrequests	01-09-2021	16-06-2023
#Uptime	01-06-2021	31-05-2023

Service name: Open Science Observatory

KPI name	From	To
#WebVisits (Unit of Access)	11-06-2021	16-06-2023
#Uptime	01-01-2021	31-05-2023

Service name: AAI

KPI name	From	To
#UserLogins (Unit of Access)	18-03-2021	16-06-2023
#APIrequests (Unit of Access)	01-01-2021	16-06-2023
#RegisteredUsers	01-01-2021	16-06-2023
#Uptime	16-06-2021	31-05-2023

6. DEVIATIONS OF VA TABLES AND KPIS

ScholeXplorer #APIRequests KPI base value (for 2019) defined in the Grant Agreement document was calculated by relying on metrics exposed directly by ScholeXplorer metrics delivery endpoint, which turned out to provide invalid values due to *Python uvicorn library* peculiarities. We could either relate to a wrongly calculated base number by relying on the same endpoint for counting current usage, or rely on a different, more trustful source of API hits. We decided to rely on log entries, which proved to be accurate. This means we cannot relate the current numbers taken from log files to the inaccurate base number calculated for 2019 from the GA document anymore and need to come up with the correct base value. According to the last 4 months of logs we still had preserved for 2019 (2019.09-2019.12) there were 185M of API requests handled by ScholeXplorer. The only way to come up with the base number for the whole year 2019 was to extrapolate this sample to the whole year, which resulted in **555M** of API hits (instead of 2Bi as originally reported in the Grant Agreement document). This is in line with the comparison of API hits served by ScholeXplorer endpoint and retrieved from the logs for the current usage in 2021 and 2022, proving that the ScholeXplorer endpoint reports 4x API requests more than it was registered in the logs.

Open Science Observatory (in-kind) does not expose publicly any API to handle requests therefore it is impossible to rely on #APIrequests as the Unit of Access. As a result of that it was replaced with #WebVisits as the UoA which is the only reasonable measure for OSO. There was no 2019 base value specified in the Grant Agreement document because service was deployed in 2020.

AAI unit of access was inconsistently defined in the Grant Agreement document: #RegisteredUsers was defined in the Units of Access summary table while #UserLogins and #APIrequests were defined in T5.7 (Authentication and Authorization Infrastructure) description. After further internal discussion it turned out #RegisteredUsers KPI was mistakenly put into the summary table, and it should not be considered as the Unit of Access. For the sake of transparency, all three KPIs are being tracked by metrics acquisition system and were reported in this document.

Some metrics (e.g. uptime for AAI, OpenAPC, OpenCitations) could not be measured since the beginning of OpenAIRE-Nexus project because OpenAIRE measurement software was not covering those services at that time, and it was implemented and deployed several months after the project started. All such cases were indicated in the “Time period coverage” section.

7. VA ASSESSMENT REPORT

1st VA External Board for VA assessment meeting outcome

The OpenAIRE-Nexus MONITOR installations were assessed in the context of Virtual Access by a board composed of international experts.

The first assessment meeting took place on 29th of November, 2022 with the following External Board members participating in the meeting:

- Andrea Manzi (EGI-ACE - EGI.eu)
- Enol Fernández (C-SCALE - EGI.eu)
- Marek Horst (OpenAIRE-Nexus - ICM UW)
- Antti Pursula (DICE - EUDAT / CSC)
- Tomasz Piontek (RELIANCE - PSNC)

The following OpenAIRE-Nexus project related deliverables:

- D1.2 - Key Performance Indicators and Analysis report
- D4.1 - PUBLISH Services: Periodic Report on VA Activities I
- D4.2 – PUBLISH Services: Periodic Report on VA Activities II
- D5.1 - MONITOR Services: Periodic Report on VA Activities I
- D5.2 - MONITOR Services: Periodic Report on VA Activities II
- D6.1 - DISCOVER Services: Periodic Report on VA Activities I
- D6.2 - DISCOVER Services: Periodic Report on VA Activities II

were shared with the VA Assessment Board one week before the VA assessment meeting.

The Virtual Access related presentation was made by Marek Horst, OpenAIRE-Nexus Virtual Access Manager. The following aspects were presented and discussed during the meeting:

- Detailed description of the Unit of access for each installation
- Description of the gathered VA indicators
    - Making sure that double accounting is avoided
- An effort to enable the automatic collection of VA metrics
    - Demo of the Dashboard based on Grafana
    - When VA Automatic Collection is not available the metrics are pushed to github using CSV files
    - MoUs as an additional condition for dashboard-focused UoA eligibility
- Logs for user access to services are kept for auditing

- Integration with EOSC VA Accounting

Additionally the following topics were raised and discussed:

- Missing or miscalculated point of reference (baseline) cases
- Retention of logs and GDPR, logs anonymization

Final assessment

All the installations under VA are offered to all the users, without selecting the researchers to whom access is provided, and free of charge what fulfils the major VA requirement. In general, the VA Assessment Board finds the implementation and operation of the VA in the OpenAIRE-Nexus project to be well managed and in line with requirements.

2nd VA External Board for VA assessment meeting outcome

The second assessment meeting took place on 1st of June, 2023 with the following External Board members and project representatives participating in the meeting:

- Charis Chatzikyriakou (C-SCALE - EODC)
- Enol Fernández (C-SCALE - EGI.eu)
- Marek Horst (OpenAIRE-Nexus - ICM UW)
- Andrea Manzi (EGI-ACE - EGI.eu)
- Raul Palma (RELIANCE - PSNC)
- Tomasz Piontek (RELIANCE - PSNC)
- Antti Pursula (DICE - EUDAT / CSC)
- Debora Testi (DICE - CINECA)

During the meeting, all the project representatives shared the gained experience in the application of the VA methods and principles for the provisioning of the different offered services. A joint report on advantages and challenges related to the VA funding model was prepared and is available as Appendix A.

8. APPENDIX A

Virtual Access assessment report

The External Advisory Board for the VA assessment of the INFRAEOSC-07 projects has been established in line with the GA requirements and in agreement with the EC and is composed by one representative from each INFRAEOSC-07 project. After a first assessment, with the aim to check and discuss and provide recommendations on the VA practices and methodologies of each project (held in December 2022), a second assessment was carried out in June 2023.

The objective of the second assessment has been to share gained experiences on the application of the VA methods and principles in the provisioning of the different offered services and to define common lessons learnt.

A plenary meeting took place online on the 1st of June 2023. During the meeting, each INFRAEOSC-07 project representative has presented the advantages and challenges in the application of the VA rules and principles to the specific service offering, and commonalities/differences were discussed.

Common lessons learnt have then been agreed and are summarised as follows.

First of all, the VA method is valued as an instrument to allow the service providers to recover the costs of used resources and services. The VA method also allows to provide the services free-at-the-point of use, even in the case of usually pay per use ones, offering the possibility to users and communities to test and start the take up of the services at no costs thus reducing the adoption barrier. It is also acknowledged that it was important to have both investment and operational costs included as part of the units cost calculation, even if not all operational costs can be claimed by all providers depending on the specific site accounting rules.

However, being the VA method new in its application for some categories of services and providers part of the INFRAEOSC-07 projects, some challenges have been identified. Overcoming these issues in future EC funded projects would make the instrument even more important and effective.

Only in few cases the VA costs were based on actual costs method only, while in most of the cases unit costs or a combination of unit and actual costs was used. The right choice of unit is thus vital to allow a meaningful representation of the services usage, but in many cases the proper unit type was difficult to define at the start of the project. At the same time, the unit costs are defined at the proposal writing phase and during the projects' execution years, many costs might change (the high increase in electricity costs of the last year is just one example). However, in the current implementation of the VA model, updates in the type and cost of units imply the need to request an amendment to the GA, which is making the process time consuming. This lack of flexibility in the VA budget management makes it also difficult to adjust the VA offering to changes in the needs from the users' communities during the projects, for example by updating the quantity of units being offered, transferring units from one installation to another, removing or adding new installations.

As already mentioned, the INFRAEOSC-07 projects collect together in their offering very diverse types of services, with different units/metrics and offered by different providers. This led to the impossibility to have a generic and automatic acquisition of the accounting data which implied in many cases to set-up manual collection of the VA accounting data; as always in manual data collection, this is time consuming and prone to errors. At the same time, the recent creation of the possibility for EOSC users to request bundles of services (from different providers but also across different projects) raised the issue on how to properly account those requests (as individual requests vs aggregated information). Another critical element related to the VA accounting is the GA requirement to maintain information needed for auditing for at least 5 years after the project ends; while logs and metrics are collected by each project, it is felt that the procedures and requirements for auditing purposes are not currently sufficiently clear (which level of details of information/logs to maintain and at which level (project level vs provider level)).

With respect to the users/communities being served, it became evident that commitment from large communities is difficult to achieve due to the short term duration of the projects and support to the VA methods. The short term VA supported offering is also perceived as discouraging for individual researchers and small groups who will not be able to pay for services via their institutions. This issue becomes particularly relevant for some categories of services like those related to data storage where the expected preservation is normally in the range of 5 to 10 years or even more. It should be also noted that having a lot of small users instead of large communities uptaking the service implies to provide much more time for users' support and engagement which might impact on the budget actual costs of the VA installations.

In summary, the INFRAEOSC-07 projects recommendations are:

- The actual cost model is in general easier to account and claim with respect to unit costs;
- More flexible procedures to update the VA offering would be advisable;
- Clearer information on the VA auditing process and requirements would be useful;
- Longer term support to the VA mechanism would reduce the barrier to adoption and make the researchers and communities more keen to update the offered services.