The JERICO-S3 project is funded by the European Commission's H2020 Framework Programme under grant agreement No. 871153. Project coordinator: Ifremer, France.

DELIVERABLE TITLE:

JERICO-S3 D11.1 First report on VA JERICO Resources access statistics and service provision

DELIVERABLE NUMBER: **D11.1**

WORK PACKAGE N° and NAME: WP11 - VA: Virtual Access

Authors: Damià Rita (SOCIB), Juan Gabriel Fernández (SOCIB), Miguel Charcos (SOCIB), Simon Keeble (Blue Lobster), Jukka Seppälä (SYKE), Peter Thijsse (Maris), Jay Pearlman (IEEE), Veronique Creach (Cefas), Antonio Novellino (EMODnet), Joaquín Tintoré (SOCIB).
With contributions from all WP11 Partners.

Involved Institutions: IFREMER, ULCO, AZTI, Cefas, CNR, CNRS-LOG, CNRS-LOV, FMI, HCMR, HZG, IH, IOW, NIVA, PdE, SMHI, SOCIB, SYKE, TALTECH, IODE, VLIZ

Due date // Submission date: 31.11.2021 // 26.11.2021

Nature: **R**
(R = Report, P = Prototype, D = Demonstrator, O = Other)

Dissemination level : **PU**
PU = Public, PP = Restricted to other programme participants (including the Commission Services), RE = Restricted to a group specified by the consortium (including the Commission Services), CO = Confidential, only for members of the consortium (including the Commission Services)

GRANT N°: 871153
PROJECT ACRONYME : JERICO-S3
PROJECT NAME : Joint European Research Infrastructure for Coastal Observatories - Science, services, sustainability
COORDINATOR : Laurent DELAUNEY - Ifremer, France - jerico@ifremer.fr

DOCUMENT TECHNICAL DESCRIPTION

Document ID	JERICO-S3-WP11-D11.1-26.11.2021-V1.0

JERICO-S3 DELIVERABLE

Joint European Research Infrastructure network for Coastal Observatory

Science, Services, Sustainability

DELIVERABLE n° WP and full title	JERICO-S3 D11.1 (WP11) First report on VA JERICO Resources access statistics and service provision
Description	First report on VA JERICO Resources access statistics and service provision
Lead beneficiary	SOCIB
Lead Authors	Damià Rita (SOCIB), Juan Gabriel Fernández (SOCIB), Miguel Charcos (SOCIB), Simon Keeble (Blue Lobster), Jukka Seppälä (SYKE), Peter Thijsse (Maris), Jay Pearlman (IEEE), Veronique Creach (Cefas), Antonio Novellino (EMODnet), Joaquín Tintoré (SOCIB)
Contributors	Susana Rubio (PdE), Julien Mader (AZTI), Laura Alonso (PdE), Lauri Laakso (FMI), Emilie Breviere (SMHI), Maria Sotiropoulou (HCMR), Kate Collingridge (CEFAS), Leonidas Perivoliotis (HCMR), Jean-Olivier Irisson (CNRS-LOV), Milla Johansson (FMI), João Vitorino (IH), Emilie Poisson Caillault (IFREMER, ULCO-LISIC), Pierre-Alexandre Hébert (IFREMER, ULCO-LISIC), Andrew King (NIVA), Luis Artigas (CNRS-LOG, ULCO), Arno Lambert (IODE), Taavi Liblik (TALTECH), Hele-Riin Pihel (TALTECH), Carlo Mantovani (CNR), Lennert Schepers (VLIZ), Lars Stemmann (CNRS-LOV), Begoña Gómez (PdE), Patrick Gorringe (SMHI), Lorenzo Corgnati (CNR), Pauline Simpson (IODE), Klas Ove Möller (Hereon), Veronique Creach (CEFAS), Patricia Cabrera (VLIZ), Paulo Nunes (IH), Gregor Rehder (IOW), Gisbert Breitbach (Hereon), Lohitzune Solabarrieta (AZTI), Yoana Voynova (Hereon), Wanda Holzhüter (IOW), Anca Hienola (FMI), Thierry Carval (IFREMER), Alan Deidun (University of Malta) and Sebastien Mancini (AODN-IMOS).
Submitted by	Damià Rita (SOCIB)

REVISION HISTORY

Revision	Date	Modification	Author
V1.0	26-Nov-2021	First Submission	Damià Rita

APPROVALS

	Name	Organisation	Date	Visa
Coordinator	Delauney Laurent	Ifremer	13-Dec-2021	LD
WP Leaders	Joaquín Tintoré & Juan Gabriel Fernández	SOCIB	26-Nov-2021	JT JF

Diffusion list

Consortium beneficiaries	Third parties	Associated Partners	other
X	X	X	

TABLE OF CONTENT

1. EXECUTIVE SUMMARY

In recent years, the EC has focused on open access to data and open science to improve the efficiency of European research. These strategies give Research Infrastructures new ways of creating impact in the scientific and general communities. Therefore, bibliometric analysis has become insufficient in assessing that impact. A complete assessment must also include the measurement of user access to research resources. Access metrics are a central pillar of these strategies and, more specifically, of this first report on VA JERICO Resources access statistics and service provision.

This deliverable D11.1 presents the actions undertaken concerning the main objectives of JERICO-S3 Virtual Access (VA), contained in WP11. These objectives are to measure access metrics to the JERICO Resources and improve those metrics with an easy and unique point of access to the resources. The following broad subjects will be presented and discussed: Virtual Access Metrics System **(VAMS)**, **Outreach** activities, VA Concept (**JERICO CORE**) and **VA Expert Panel**. The activities carried out during the initial 19 months of the project are described.

A central system to monitor the access to each VA service has been created. It is called the Virtual Access Metrics System (**VAMS**). For each VA infrastructure detailed in WP11, VAMS contains a database with information on how and when the infrastructure is accessed and one dashboard illustrating this information. The system collects metrics from a diverse range of data sources. The two main types of data sources are log files and Google Analytics API. For most of the infrastructures, VAMS collects metrics automatically. For those infrastructures that could not (technically or legally) provide direct access to their metrics for an automatic collection, Google Forms for manual reporting, have been set up. In that case, VAMS automatically synchronizes with the submitted form responses. The system will be described functionally and technically in section 3.1.4. Each one of the aforementioned dashboards is reproduced in Annexe 1. JERICO-S3 coordinators, the representatives of each infrastructure and members of the VA Experts Panel will have online access to the up-to-date dashboards.

Outreach and communication activities are important means of increasing the access metrics of a VA infrastructure. Each VA infrastructure was responsible for reporting the outreach activities they have performed during the period this report covers. The reported outreach activities can be found in Annexe 2.

JERICO-S3 proposes a new approach to improve the access to VA services and increase their access metrics: **JERICO CORE**. JERICO CORE, formerly known as e-JERICO, is under development and will be a central Virtual Infrastructure whose role will be to facilitate the findability of coastal oceanographic resources, including VA. Accordingly, a framework has been created to present a vision for VA of JERICO called the VA Concept. In this VA Concept, the role of JERICO CORE is to store the metadata of resources and to derive the relationships among them. The metadata of a VA service contains the necessary information to help the user access the data directly from the VA service itself. JERICO CORE also hosts

additional services that have been specifically developed, called Data-to-Products Thematic Services (D2PTS).

Finally, a **Virtual Access Experts Panel** has been formed. Its role is to evaluate the VA services described in this report. A second deliverable (D11.2) will follow , and will contain the feedback provided by the panel. In 2023, D11.3 will be composed with a similar type of structure and will update the metrics data. D11.4 will follow with feedback from the Virtual Access Expert Panel. The panel members are Anca Hienola, Antonio Novellino, Thierry Carval, Simon Keeble, Alan Deidun and Sebastien Mancini. More information about the Panel is found in section 3.4.

2. INTRODUCTION

Research Infrastructures (RIs) are facilities, resources and services that are used by the research communities to conduct research and foster innovation in their fields. Where relevant, they may be used beyond research, for example for education or public services. They include major scientific equipment or sets of instruments; knowledge-based resources such as collections, archives or scientific data; e-infrastructures such as data and computing systems and communication networks; and any other infrastructure of a unique nature essential to achieving excellence in research and innovation. Such infrastructures may be 'single-sited', 'virtual' or **'distributed'**.[1]

A **distributed RI** consists of a Central Hub and interlinked National Nodes. According to ESFRI's Glossary[2], RIs need, among other things, to:

- Have a common access policy and provide for a single point of access for all users with a support structure dedicated to optimising the access for the proposed research; - have a user programme designed to absorb a considerable part of the total capacity of the RI.
- Identify and agree upon relevant and measurable Key Performance Indicators (KPI) addressing both excellence of scientific services and sustainability of operation
- Define a joint investment strategy aimed at strengthening the RI through the Nodes and common/shared facilities.

JERICO-S3 is a project funded by the European Commission's H2020 Framework Programme. The ambition of its members is the creation of an integrated pan-European multidisciplinary and multi-platform **distributed research infrastructure** dedicated to a holistic appraisal of coastal marine system changes.

In order to build on previous research results, encourage collaboration and avoid duplication of effort, speed up innovation and involve citizens and society, the H2020 Framework Programme encourages the provision of **open access** to research data[3]. Moreover, since 2016 the European Commission (EC) has enlarged the focus to **open science**. Open science brings changes to each step in the scientific process, such as the emergence of alternative systems to establish scientific reputation, changes in the way the quality and

[1] Article 2 (6) of the Regulation (EU) No 1291/2013 of 11 December 2013: "Establishing Horizon 2020 - the Framework Programme for Research and Innovation (2014- 2020)"

[2] Adapted from ESFRI's Glossary: https://www.esfri.eu/glossary

[3] Article 18 (2) of the Regulation (EU) No 1291/2013 of 11 December 2013: "Establishing Horizon 2020 - the Framework Programme for Research and Innovation (2014- 2020)"

impact of research are evaluated, the growing use of scientific blogs, open annotation and open access to data and publications.[4]

In 2016, the '**FAIR** Guiding Principles for scientific data management and stewardship' were published[5] in *Scientific Data*. The authors intended to provide guidelines to improve the **Findability, Accessibility, Interoperability, and Reuse** of digital assets. The principles emphasise machine-actionability (i.e., the capacity of computational systems to find, access, interoperate, and reuse data with none or minimal human intervention) because humans increasingly rely on computational support to deal with data as a result of the increase in volume, complexity, and creation speed of data.[6]

Virtual Access (VA) is the financial instrument to reimburse the access provisioning costs to access providers. The European Commission provides this instrument to increase the sharing of research infrastructures and services that otherwise would not be available to international user groups. In VA, the **services** or resources – also called "installations" – have to be made available 'free of charge at the point of use'. VA must be free and open access to services through communication networks to resources needed for research, without selecting the researchers to whom access is provided.[7]

In a world of open science, where the performance of researchers or research institutions cannot be measured by bibliometric data only, access metrics to a VA service are important additional indicators. As a matter of fact, the OECD has developed a set of metrics[8] to assess the overall impact of research infrastructures. Some of them are related to data access: S6-Number of scientific users, T27-Data sharing (Number of data requests, Number of data accesses).

Additionally, the **Leiden Manifesto**[9] in 2015 or the DORA Declaration in 2012[10], among other recent initiatives, proposed that the research performance should not be delegated to

[4] "Open Innovation, Open Science, Open to the World – a vision for Europe" European Commission, 2016 (doi:10.2777/061652)

[5] Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016). https://doi.org/10.1038/sdata.2016.18

[6] Extracted from https://www.go-fair.org/fair-principles/

[7] Adapted from EOSC-hub "D13.1 Periodical assessment of the services" https://documents.egi.eu/public/ShowDocument?docid=3409

[8] "Reference framework for assessing the scientific and socio-economic impact of research infrastructures" OECD (https://doi.org/10.1787/3ffee43b-en)

[9] Hicks, D., Wouters, P., Waltman, L. et al. Bibliometrics: The Leiden Manifesto for research metrics. Nature 520, 429–431 (2015). https://doi.org/10.1038/520429a

[10] https://sfdora.org

one single metric, but based on the evaluation of the research activities themselves with the support of a set of metrics that are constantly reviewed and adapted to the circumstances. These alternative metrics are known as altmetrics.

Access metrics are quantitative assessments used for assessing user access to research resources and provide a first insight into their level of dissemination. The role of access metrics is not only to measure the research performance as they also allow researchers to assess the **fitness for use** of their results. Thus, closing a feedback loop that allows researchers to make decisions to improve their dissemination to the needs of the data users.

Virtual access metrics are the access metrics to **VA services** (ie. a VA point of access with specific purpose and technology). One institution can host a variety of VA services. In this document, the VA services of an institution and the team that manages them will be referred to as a **VA Infrastructure**. In JERICO-S3, a list of already-existing VA infrastructures was selected to receive financial support under VA. This list is provided in section 3.1.5 of this document.

Access metrics to the VA services of JERICO-S3 WP11 are a central pillar of this document. However, this is one more step in the path JERICO has been walking previously. During JERICO-NEXT[11] (2015-2019), the availability indicators (AI) were defined based on previous work by EMODnet Mediterranean CheckPoint and VA services were monitored using the AIs and access metrics.

This document explains **VAMS**, an automatic system that collects access metrics of VA services and shows them in a centralised place (section 3.1.4). The access metrics **dashboard** that has been generated for each VA service is described (Annexe 1). The **outreach activities** each VA infrastructure has carried out are presented (Annexe 2). The **VA Concept** is detailed, which describes the role to be played in the future of JERICO's VA (Section 3.3) by e-JERICO, from now on referred to as JERICO CORE (Coastal Ocean Resource Environment), aligned with the long term development of a European coastal ocean infrastructure, associated with the Global Coast Predict Programme (endorsed by the UN Decade for Sustainable Oceans), and also aligned with JERICO-DS activities and vision. To conclude, the **Virtual Access Expert Panel** and its role is introduced (Section 3.4).

[11] https://www.jerico-ri.eu/previous-project/jerico-next/

3. MAIN REPORT

This report corresponds to the 19 first months of JERICO-S3: from February 2020 to August 2021.

The report starts by describing the Virtual Access Metrics System and its results. Secondly, the outreach activities carried out by each VA provider included in WP11 are reported. Next, the VA Concept, a framework that defines the role of JERICO CORE, is introduced. Finally, the Virtual Access Expert Panel is presented.

3.1. Virtual Access Metrics

3.1.1. Need for Metrics

Access metrics to VA services are an essential piece of information to manage a research infrastructure (both at the strategic and planning levels). They allow managers to make informed decisions and prioritise actions in a data-driven way. For example, a new sensor could be bought and installed due to the specific interest in a certain type of dataset or an outreach campaign could be launched at a specific time of the year to target the right people at the right moment.

3.1.2. Previous Work - We Stand on the Shoulders of Giants

JERICO-S3 is not the first JERICO project to assess VA services. The efforts carried out under WP6 and WP8 of JERICO-NEXT are very relevant. Specifically, a lot of insight has been gained from D5.16[12] (definition of the availability indicators used in WP6) and D8.12[13] (definition of the questionnaire for the periodic assessment of VA services) of JERICO-NEXT. Additionally, both JERICO-NEXT and this document have benefited a lot from the work carried out by Med Sea Checkpoint: specifically, their Data Adequacy Report[14] .

The following factors, however, have supported the need for an alternative approach:

[12] Linking JERICO-NEXT activities to a Virtual Access infrastructure (https://www.jerico-ri.eu/download/jerico-next-deliverables/JERICO-NEXT-Deliverable-5.16_V2.pdf)

[13]Template for assessment of web-based portals under JERICO-NEXT WP6 (https://www.jerico-ri.eu/download/jerico-next-deliverables/JERICO-NEXT-Deliverable_8.12_V4.pdf)

[14]Growth and Innovation in Ocean Economy - Gaps and Priorities in Sea Basin Observation and Data (https://emodnet.ec.europa.eu/sites/emodnet.ec.europa.eu/files/public/D11.2-revised-V11.pdf)

- The concept of VA is broader in JERICO-S3. VA in JERICO-NEXT included data and data products. In JERICO-S3, however, platforms, software, e-training and best practices documents are also included. This makes it impossible to develop a questionnaire that is relevant for all.
- JERICO CORE is under development. When JERICO CORE is operational, it will provide a central point of search of JERICO resources (VA services will be some of the resources contained by the catalogue). This will improve the findability of the resources as they will be catalogued by a system that understands them. Hence, the measurement of visibility indicators of each VA service will become less relevant.
- JERICO-S3 Task 6.4.2 will be evaluating, using the GOFAIR metrics[15], the FAIRness of data and facilitating the flow of data to aggregators

Therefore, in this deliverable, the Virtual Access Metrics System (VAMS) is introduced. It consists of a system that automatically and periodically collects access metrics from the different VA services and is capable of analysing and presenting them in multiple ways. An additional advantage is that the effort of creating the system is done only once regardless of the number of reports that need to be delivered.

3.1.3. What is being measured?

To monitor access metrics to VA services, it is possible to consider three different perspectives depending on the questions that need to be answered:

1. Focus on IT. In this case, the question is: "Is the IT infrastructure able to cope with the existing requests?". To answer it, all traffic needs to be counted, including traffic from robots, to invalid URLs...
2. Focus on the VA service. In this case, the questions are: "What is the usage of my VA service?" and "What resource is being accessed?". To answer these questions, certain traffic needs to be filtered out (i.e: from robots to invalid URLs…) to effectively analyse the requests.
3. Focus on the asset[16]. In this case, the question is: "What is the impact of the content I produce from the point of view of virtual access?" To answer this question, institutions should be able to monitor the assets they generate wherever they are accessed (i.e: data aggregators, derived (data)products, scientific publications…)

[15] A design framework and exemplar metrics for FAIRness
Mark D. Wilkinson, Susanna-Assunta Sansone, Erik Schultes, Peter Doorn, Luiz Olavo Bonino da Silva Santos, Michel Dumontier
bioRxiv 225490; doi: https://doi.org/10.1101/225490
Now published in Scientific Data doi: 10.1038/sdata.2018.118

[16] In this section, "asset" is to be understood as "piece of content accessible via VA". One asset could have different points of access and the access metrics of an asset integrate the access metrics in each point of access.

The access metrics provided in this deliverable are based on the second perspective. But they have the finest granularity available to provide as much detail as possible about the assets that are being accessed.

It has not been considered the first perspective to be relevant in the context of this document. The third perspective is very interesting. However, currently, it is not possible to systematically connect the access metrics of one asset at different access points. In the next section, the issue is analysed in depth.

3.1.3.1. Is this FAIR?

As previously mentioned, it would be very interesting for an institution to use access metrics to assess the overall impact of the assets it produces.

The main objective of showcasing assets under Virtual Access is maximising their distribution and their impact on society. One way of achieving this is by adhering to the principles of Findability, Accessibility, Interoperability and Reusability (FAIR). In consequence, assets can be downloaded once and used many times in many different ways. This is particularly clear in the case of software packages (they may be downloaded once and executed several times).

While these principles are great to make scientific and technical efforts as useful as possible, they make it very hard to measure their overall impact. For instance, an institution could have produced a dataset that receives very little traffic at their VA service that, on the other hand, might be critical for another organisation to create a key product in their catalogue. There is currently no way for the producer of the dataset to automatically match its asset to the other product or to receive access metrics from it in a standard way. However, this information is very relevant to assess the impact of the dataset.

Finding a solution for this is a long-term complex project. But FAIRness is the first step, as the use of metadata helps to track the provenance of datasets. Having precise ways to communicate access metrics (including what units they are measured in, how they are calculated, what filters are applied...) among teams will probably be another important part of the solution. This could be accomplished either by imposing standards on the access metrics or by attaching metadata to them. Finally, a way is needed to discover all the resources related to the one we want to assess. JERICO CORE's catalogue (see section 3.3 for more details) could be a suitable tool for this.

3.1.3.2. Apples and Oranges

Metrics offered by VAMS cannot be taken as absolute and they cannot be compared across VA services. This is because of the different nature of the VA services, the content they

expose and the users that access them. For example, the access metrics of a VA service that is accessed by many final users cannot be compared with the access metrics of a VA service whose users are, for the most part, aggregators that distribute the content. Even different users of a given VA service could have very different significances while generating the same metrics: it is not the same a user that casually browses a website than a member of a national forecasting service who is getting relevant information for their work. Additionally, different VA services have had different lengths of time to grow their user base and will, thus, have different metrics because of their different ages.

Instead, it is especially interesting to see the trends that develop over time and interpret them based on the context of the VA (Did an aggregator stop getting data directly and is getting it via a third party? Did COVID make it impossible to collect data of this type and thus traffic is reduced?...). Trends can be used to find out the relevant use cases, to deduce the factors that affect the volume of access and to put in place observation strategies to mitigate the detected weaknesses.

3.1.4. VAMS Architecture

The VA Metrics System (VAMS) is a system that has been developed to monitor the access metrics of the VA services under WP11. Its purpose is to provide a central point where the access metrics can be viewed.

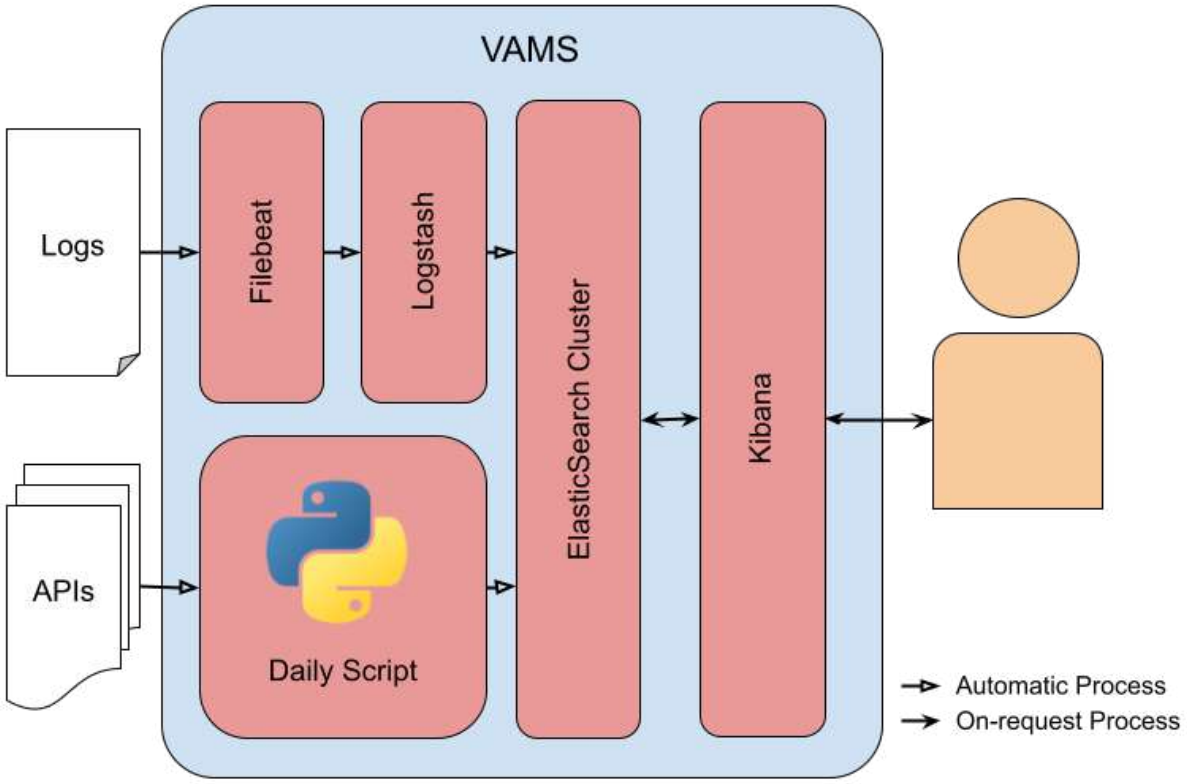
To implement VAMS, the Elastic stack[17] was chosen. Elastic is a popular, open-source project that centres around a fast, distributed database called ElasticSearch. The distributed nature of the database, allows it to scale across different servers if the amount of data requires extra hardware to handle. Elasticsearch is a no-SQL database that stores schema-free JSON documents and it uses an HTTP interface. The Elastic stack also includes, among others:

- Kibana: a web application that specialises in managing ElasticSearch databases and creating reports of the data they contain.
- Logstash: a very flexible microservice that takes a series of events and modifies the data of each event according to a configuration. It can remove unnecessary or private data, it can discard events and it can transform and enrich the data in the original event (provide approximate geolocations, categorise URLs…)
- Filebeat: a piece of software whose job is to monitor files and detect new lines each time they get added. It can perform basic filtering of the data before transmitting it over time to Logstash or ElasticSearch.

[17] https://www.elastic.co/

The main part of VAMS consists of two nodes of Elasticsearch that contain the information. One node of Kibana presents that information.

Depending on the technology used by each VA service, VAMS has a different way of accessing metrics. In other words, different integrations have been developed for different VA services. The next diagram shows these different integrations.



Overview of VAMS architecture

In the case of VA services whose metrics data source consisted of log files, Filebeat has been used to monitor the relevant files and detect new lines. Filebeat transfers each line to Logstash which, after processing it, sends it to ElasticSearch for storage. To estimate the number of users, the number of distinct IPs is used. This is an estimation, as many people can share an IP (thus, underestimating the number of users) and one's IP can change over time (thus, overestimating the number of users).

Alternatively, for VA services whose metrics are accessible via an API, a Python script handles the data. The script connects to the relevant API daily, collects and parses the necessary data and sends it to ElasticSearch for storage. The list of APIs that are being handled by VAMS is:

- Google Analytics Reporting API v4[18]: used for VA services that consist of a website that is monitored via Google Analytics
- CRAN-logs[19]: used to monitor the downloads of software packages written in R available in the CRAN archive
- Custom CSV endpoints: some VA providers decided to expose a CSV file with the relevant metrics at an agreed-upon URL.
- ElasticSearch API[20]: used to synchronise VAMS with an external ElasticSearch database for VA providers who are already collecting and analyzing their access statistics using similar technology to VAMS.
- Github Traffic API[21]: used to monitor software packages that are stored on Github. Unfortunately, Github only saves the access metrics for 15 days. So, access data previous to 15 days before the development of the connection is lost.
- ICOS stats API: The calls done to the stats backend done by https://data.icos-cp.eu/stats/ have been reverse-engineered to obtain access metrics of datasets that are published via SOCAT. The support provided by the team at SOCAT deserves to be acknowledged here. They went so far as to upgrade their API to accommodate VAMS.
- SQL: used to monitor accesses to systems that store event data in a relational database.

Finally, for VA providers that could not integrate automatically into VAMS (for legal or technical reasons) a Google Form was created for periodic reporting. Data submitted via these forms is stored in Google Spreadsheets and the Google Spreadsheet API[22] is used to synchronise VAMS daily.

3.1.5. VA Infrastructures Integrated into VAMS

As mentioned above, different VA services had different types of integration. The following chart describes the various types of methods of access to statistics across the VA network.

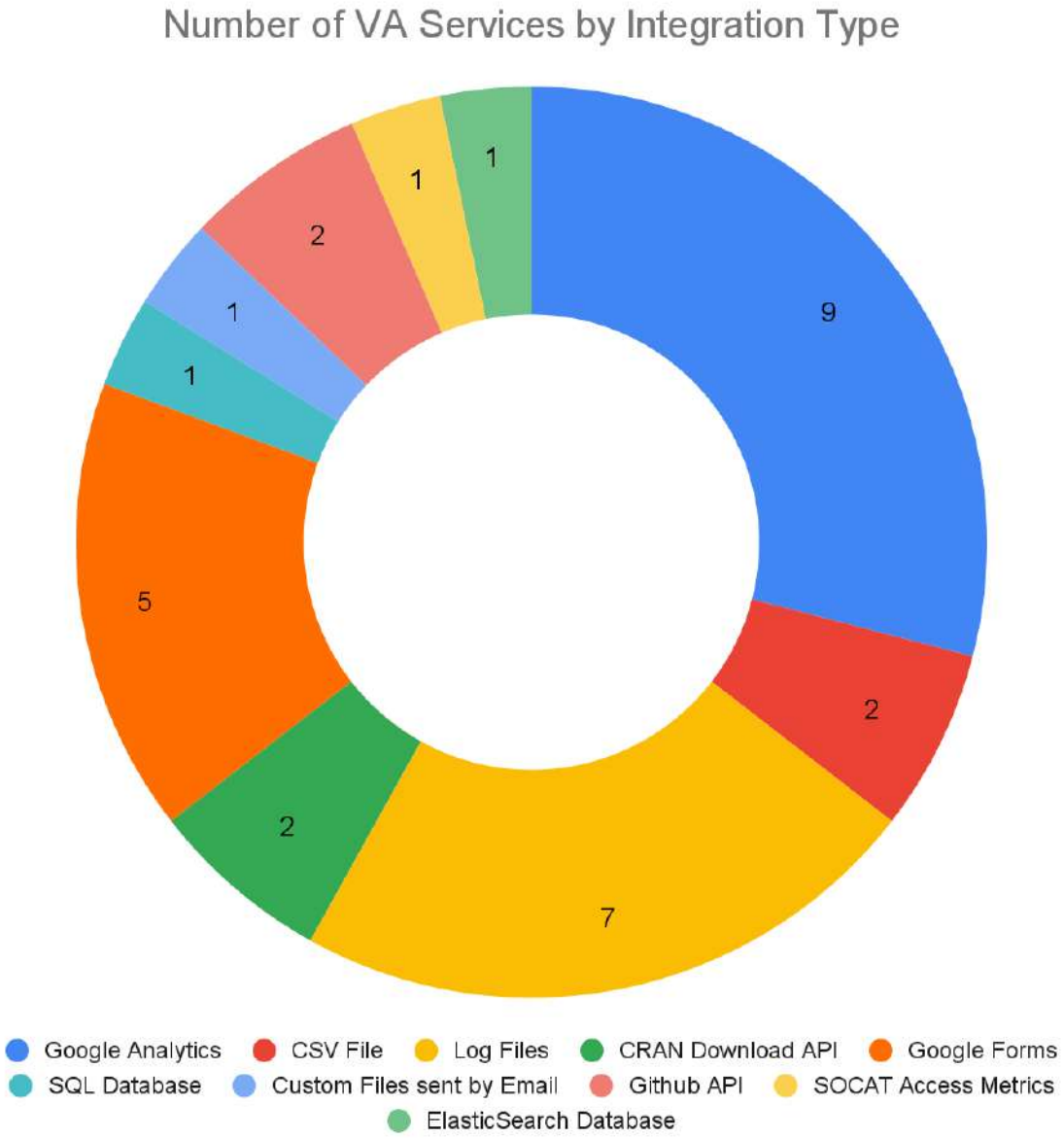
[18] https://developers.google.com/analytics/devguides/reporting/core/v4

[19] https://cranlogs.r-pkg.org

[20] https://www.elastic.co/guide/en/elasticsearch/reference/current/rest-apis.html

[21] https://docs.github.com/en/rest/reference/repos#traffic

[22] https://developers.google.com/sheets/api/reference/rest



The following is a short technical description of each VA infrastructure included in JERICO-S3 WP11.

3.1.5.1. VA Infrastructures under Development

When the JERICO CORE application development is completed, IFREMER and SOCIB will collaborate to host and maintain it respectively over the long term. Therefore, no access metrics have been generated so far.

ID	1.1

Name	DATARMOR
Institution	IFREMER
Short Description	Datarmor is an IT Infrastructure dedicated to Marine Sciences and located in Brest-France and funded in the framework of the State-Region Plan by the European Commission (FEDER funds), the French Government, Britany Region and local authorities in association with Ifremer and SHOM. Datarmor will host the JERICO e-Infrastructure during the project, including Virtual Research Environment and the associated data when High-Performance Computing and/or powerful access to data will be required.

ID	33.2
Name	JERICO RI e-Infrastructure
Institution	SOCIB
Short Description	VA framework that provides access to the most important JERICO Resources (Catalogued) and Pilot Data-to-Products Thematic Services (D2PTS).

3.1.5.2. Active VA Infrastructures

The following are the VA infrastructures that are active at the time of publishing this document. For each, a description of its integration with VAMS and the annexe that contains the actual metrics is included.

The structure of each annexe depends on the type of integration used by VAMS to receive access data. In some cases, since there was a better knowledge of the VA service, it has been possible to produce more metrics from the access data (i.e: it has been possible to extract information about what was being accessed via the URL).

ID	1.2
Name	Mawenzi
Institution	IFREMER, ULCO-LISIC
Short Description	Software packages written in R to help scientists, as well as stakeholders, in the modeling and interpretation of data (time series data from gliders/ferrybox or cytometry data), with some tutorials and user guidelines.

VA Services	The VA Infrastructure consists of the following R packages: DTWBI, DTWUMI and uHMMweb.
Monitoring Method	To monitor the access to the packages, it was decided to monitor the number of downloads from the CRAN pages and the visits to the documentation webpages[23].
Metrics	Mawenzi metrics can be found in Annexe 1.1.

ID	4.1
Name	EU HFR Node/AZTI
Institution	AZTI
Short Description	Open access to different softwares for processing and analysing coastal data. The tools aim to support a sustainable management of the ocean and its resources.
VA Services	The VA Infrastructure consists of a Github repository[24].
Monitoring Method	The Github API has been used to monitor the number of visits to the GitHub repository
Metrics	EU HFR Node/AZTI metrics can be found in Annexe 1.2.

ID	6.1
Name	CefMAT
Institution	Cefas
Short Description	Access to marine environmental assessments to address policy questions. Users can efficiently produce high quality visualisations and summary statistics tailored to their queries. It also gives access to specific products created under CMEMS and visualisation of aggregated DOI biological datasets from Cefas Data Hub.

[23] http://mawenzi.univ-littoral.fr/DTWBI/ , http://mawenzi.univ-littoral.fr/DTWUMI/ and http://mawenzi.univ-littoral.fr/uHMMweb/

[24] https://github.com/Fundacion-AZTI/JRadar

VA Services	The VA Infrastructure consists of a website[25].
Monitoring Method	The website is monitored using Google Analytics. Data is sent to VAMS by monthly filling a Google Form
Metrics	CefMAT metrics can be found in Annexe 1.3.

ID	7.1
Name	CNR TirLig
Institution	CNR
Short Description	Provide access and distribution of sea surface current data derived from High Frequency Radar acquisitions.
VA Services	The VA Infrastructure consists of a THREDDS data server[26] containing HF Radar data of the Ligurian coast near La Spezia and Cinque Terre and a website[27] with visualisations of that data and links to the mentioned THREDDS data server.
Monitoring Method	The THREDDS data server is monitored using the log files and the website is monitored with Google Analytics via API.
Metrics	CNR TirLig e-infrastructure metrics can be found in Annexe 1.4.

ID	7.2
Name	EU HFR Node/CNR
Institution	CNR
Short Description	The EU (European) HFR (High Frequency Radar) Node/CNR is the contribution of CNR to the EU HFR Node Competence Center for HFradar data management.

[25] https://www.cefmat.org

[26] http://150.145.136.27:8080/thredds/HF_RADAR/TirLig/TirLig_catalog.html

[27] http://radarhf.ismar.cnr.it/

VA Services	The VA Infrastructure consists of a software package written in MATLAB [28] and a remote processing service[29] that uses the software to process HF Radar files.
Monitoring Method	The software package is monitored using the views at the GitHub repository and the processing service is monitored using the event data stored in a SQL
Metrics	EU HFR Node/CNR metrics can be found in Annexe 1.5.

ID	8.1
Name	CytoFluoTool
Institution	CNRS-LOG, ULCO
Short Description	Automated techniques, and complementary to automated image analysis, automated flow cytometry (FCM) and multispectral fluorometry (MSF).
VA Services	The VA Infrastructure consists of a software package written in R (RclusTool)
Monitoring Method	To monitor the access to the VA service, it has been decided to monitor the number of downloads from CRAN and the visits to their documentation webpage[30]
Metrics	CytoFluoTool metrics can be found in Annexe 1.6.

ID	8.2
Name	Ecotaxa
Institution	CNRS-LOV
Short Description	Web application that provides services for users to handle large image datasets and their associated metadata. System suited for imaging

[28] https://github.com/LorenzoCorgnati/HFR_Node__Historical_Data_Processing

[29] http://150.145.136.36/

[30] http://mawenzi.univ-littoral.fr/RclusTool/

	sensors such as UVP5, ZOOSCAN, ZOOCAM, FLOWCAM, IFCB, microscopic imaging, and confocal microscopy, that are used in marine sciences.
VA Services	The VA Infrastructure consists of a website[31].
Monitoring Method	EcoTaxa has its own metrics system that synchronizes to VAMS via a custom CSV file
Metrics	EcoTaxa metrics can be found in Annexe 1.7.
Funders	• OCEANOMICS[32] • Partner University Fund[33] • CNRS LEFE program[34] • Belmont Forum WWW.PIC project[35] • H2020 Blue-Cloud project[36] • JERICO-S3[37]

ID	15.1
Name	Utö Atmospheric and Marine Research Station
Institution	FMI
Short Description	Software for instrument-computer interfaces, automated warning message system, data transmission systems, basic QC processes, data visualisation tools and a www-page open for public use.

[31] https://ecotaxa.obs-vlfr.fr/

[32] http://www.oceanomics.eu/

[33] https://face-foundation.org/higher-education/partner-university-fund/

[34] https://programmes.insu.cnrs.fr/lefe/

[35] https://www.belmontforum.org/projects/world-wide-web-of-plankton-image-curation/

[36] https://www.blue-cloud.org

[37] https://www.jerico-ri.eu/

VA Services	The VA Infrastructure consists of a website[38].
Monitoring Method	The website is monitored using Google Analytics API.
Metrics	Utö metrics can be found in Annexe 1.8.

ID	16.1
Name	POSEIDON Multi platform observatory Data Center
Institution	HCMR
Short Description	Integrated observatory located in the Eastern Mediterranean, which has adopted a multiplatform-multiparameter approach with the current system's status including open and coastal sea fixed platforms, deep-ocean observatories, a Ferrybox system, glider missions and Argo profiling floats.
VA Services	The VA Infrastructure consists of a website[39] and an API[40].
Monitoring Method	The website is monitored via Google Analytics and VAMS gets the metrics via Google Analytics API. The API is monitored using the log files.
Metrics	POSEIDON metrics can be found in Annexe 1.9.

ID	17.1
Name	COSYNA
Institution	Hereon
Short Description	Monitor real-time conditions and provide short-term forecasts, data, and data products to help assess the impact of anthropogenically induced change. It includes data from Hereon HF Radar, FerryBoxes, Underwater Nodes and operational models (Circulation, Waves) as well as models with assimilated data.

[38] http://swell.fmi.fi/Uto/

[39] https://poseidon.hcmr.gr/

[40] https://api.poseidon.hcmr.gr/swagger/

VA Services	The VA Infrastructure consists of a website[41]
Monitoring Method	The website is monitored using a custom system described in the paper "Accessing Diverse Data Comprehensively – CODM the COSYNA Data Portal"[42]. This system requires users to log in. So, downloads from scrappers, which used to be very important, are not included in the access metrics. Metrics are sent to VAMS by monthly filling a Google Form.
Metrics	COSYNA metrics can be found in Annexe 1.10.

ID	19.1
Name	HIDROGRAFICO+
Institution	IH
Short Description	Access to data collected by the real-time monitoring infrastructure for the Portuguese waters (MONIZEE system) that is operated by Instituto Hidrografico.
VA Services	The VA Infrastructure consists of a website[43] which presents data provided by a webservice.
Monitoring Method	The webservice has its own metrics system that collects the data in its own ElasticSearch database. VAMS collects the new metrics daily from the ElasticSearch API. The website is monitored using Google Analytics API.
Metrics	HIDROGRAFICO+ metrics can be found in Annexe 1.11.

ID	21.1
Name	VOS Finnmaid GHG - BGC
Institution	IOW

[41] https://www.hereon.de/institutes/coastal_ocean_dynamics/cosyna/index.php.en

[42] DOI: 10.5194/os-2016-6 (http://dx.doi.org/10.5194/os-2016-6)

[43]https://geomar.hidrografico.pt/

Short Description	Measurement of trace gases on the ferry Travemünde/Germany to Helsinki/Finland.
VA Services	The data is distributed via ICOS (SOCAT[44] database)
Monitoring Method	Access metrics are obtained by VAMS via the API of the Carbon Portal's Statistics page[45] which ICOS' tech team gracefully upgraded to accommodate VAMS's needs.
Metrics	VOS Finnmaid GHG - BGC metrics can be found in Annexe 1.12.

ID	25.1
Name	NorFerry/NorSOOP
Institution	NIVA
Short Description	Coastal observing data from FerryBoxes, descriptions of data types, and ocean literacy educational stories are provided for public interaction in the form of 24 inch touchscreen display consoles on passenger vessels that are updated pseudo-real-time.
VA Services	The VA Infrastructure consists of the application available on the touchscreens.
Monitoring Method	Each touchscreen collects access metrics and stores them on a server. Periodically, NIVA staff collects this data and transmits it to VAMS
Metrics	NorFerry/NorSOOP metrics are partial due to COVID shutdowns of passenger vessels, but they can be found in Annexe 1.13.

ID	28.1
Name	PORTUS observing and forecasting system
Institution	PdE
Short Description	Oceano-meteorological network and forecasting system that integrates different observing platforms (buoys, tide gauges and HF-radars) with numerical models along the Spanish coast. All data and derived products

[44] https://www.socat.info/

[45] https://data.icos-cp.eu/stats/

	are distributed through PORTUS early warning system and visualisation tool. Additionally, data from the numerical models, the HF radars and the tide gauges are presently available in the PORTUS OPeNDAPsystem.
VA Services	The VA Infrastructure consists of a website[46] and a THREDDS data server[47].
Monitoring Method	The THREDDS data server is monitored using log files.
Metrics	PORTUS metrics can be found in Annexe 1.14.

ID	32.1
Name	Swedish Oceanographic Data Centre - Toolboxes in marine data management
Institution	SMHI
Short Description	Access to Swedish oceanographic data and with the use of toolboxes; handling, quality control and analyses of such data.
VA Services	The VA Infrastructure consists of 3 websites (SMHI general website[48], OpenData View[49] and Shark Web[50]) and 2 APIs (OpenData[51] and SharkData[52]).
Monitoring Method	Each has its own metrics system. But all metrics are sent to VAMS via a monthly Google Form submission.
Metrics	The relevant metrics can be found in Annexe 1.15.

[46] http://portus.puertos.es

[47] http://opendap.puertos.es

[48] https://www.smhi.se

[49] https://opendata-view.smhi.se

[50] https://sharkweb.smhi.se

[51] https://opendata.smhi.se

[52] https://sharkdata.smhi.se

ID	33.1
Name	SOCIB Data Centre Multi-Platform Observatory
Institution	SOCIB
Short Description	Access to the multi-platform observing system of the Balearic Islands in the western Mediterranean.
VA Services	The VA Infrastructure consists of a THREDDS data server[53], a data API[54] and a data catalogue[55].
Monitoring Method	The THREDDS data server and the API are monitored using log files, whereas the data catalogue is monitored using Google Analytics and its AP
Metrics	SOCIB metrics can be found in Annexe 1.16.

ID	34.1
Name	SYKE-ALG@LINE
Institution	SYKE
Short Description	Monitors the state of the Baltic Sea using ferrybox systems. Collected data includes simultaneous measurements of physics, biogeochemistry and biology carried out in two ferrylines.
VA Services	The VA Infrastructure consists of a set of web pages [56] with information about Alg@line and its data
Monitoring Method	They are monitored using Google Analytics and the metrics are sent to VAMS using Google Forms.

[53] https://thredds.socib.es

[54] http://api.socib.es

[55] http://apps.socib.es

[56] https://www.marinefinland.fi/en-US/The_Baltic_Sea_now/Automatic_observations_from_ships, https://www.marinefinland.fi/en-US/The_Baltic_Sea_now/Algal_bloom_observations, http://swell.fmi.fi/Algaline/ and https://www.finmari-infrastructure.fi/ferrybox/

Metrics	Alg@line metrics can be found in Annexe 1.17.

ID	35.1
Name	Keri Island research station
Institution	TALTECH
Short Description	Part of the Network of experimental research stations of the Estonian Observatory that measures high-frequency full depth profiles of temperature, salinity, turbidity, oxygen content, chlorophyll-a and phycocyanin fluorescence in the Central Gulf of Finland.
VA Services	The VA Infrastructure consists of a webpage[57]
Monitoring Method	It is monitored using Google Analytics and the metrics are sent to VAMS using Google Forms
Metrics	Keri Island metrics can be found in Annexe 1.18.

ID	36.1
Name	OBPS-OTGA
Institution	IODE of UNESCO-IOC
Short Description	**Ocean Best Practices System** (OBPS): Open access, permanent, digital repository of community best practices in ocean-related sciences maintained by the IODE of the UNESCO-IOC as an IOC coordinated activity **The OceanTeacher Global Academy** (OTGA) Project aims at building equitable capacity related to ocean research, observations and services in all IOC Member States. **AquaDocs**: Open access thematic document repository covering the natural marine, coastal, estuarine/brackish and freshwater environments maintained by the UNESCO/IOC InternationaL Oceanographic Data and Information Exchange (IODE) and the International Association of Aquatic and Marine Science Libraries and Information Centers (IAMSLIC) with support from the FAO Aquatic Sciences and Fisheries Abstracts.

[57] https://taltech.ee/meresusteemide-instituut/mereinfo

VA Services	The VA Infrastructure consists of three different websites: AquaDocs[58] (an oceanographic documentation open-access repository), OceanTeacher[59] (an oceanographic web-based training platform) and OceanBestPractices[60] (a best practices repository).
Monitoring Method	The websites are monitored via Google Analytics and VAMS is synchronised using Google Analytics API.
Metrics	The relevant metrics can be found in Annexe 1.19.

ID	38.1
Name	VLIZ Marine Data Archive
Institution	VLIZ
Short Description	The infrastructure is an archival platform that provides a java based interface with a connected MS SQL Server database for upload and documentation of data files.
VA Services	The VA Infrastructure consists of a website[61]
Monitoring Method	It has its own custom metrics system and VAMS is synchronised via a CSV file that VLIZ generates for this purpose
Metrics	The relevant metrics can be found in Annexe 1.20.

3.2. VA Outreach

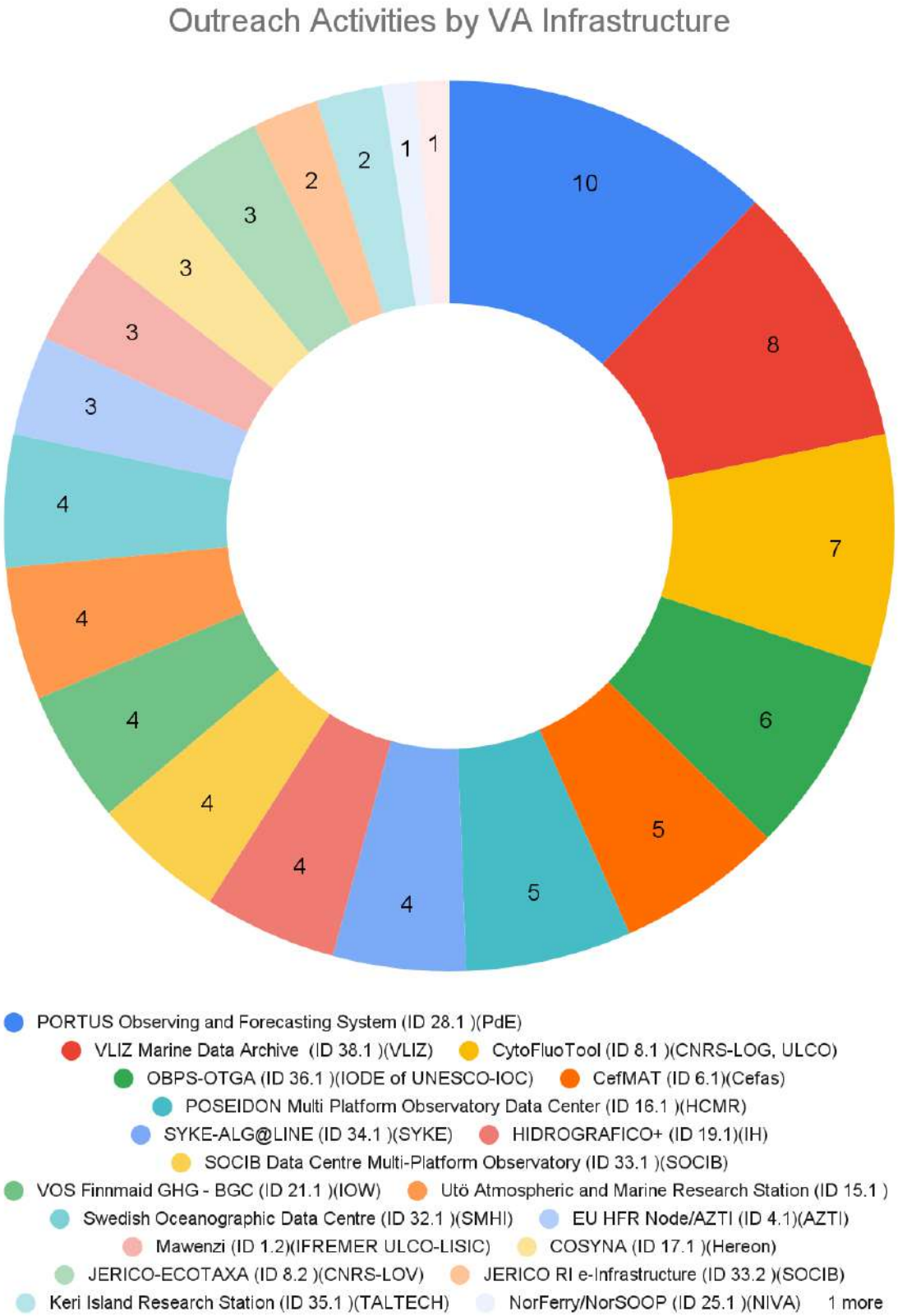
One tool to increase the access metrics to one's VA infrastructure is to engage in outreach activities that promote it. One of the responsibilities of VA providers is to regularly reach out to new audiences to increase the user base. What follows is a brief analysis of these activities. In Annexe 2, all the outreach activities carried out by partners of JERICO-S3 WP11 related to their VA services will be reported.

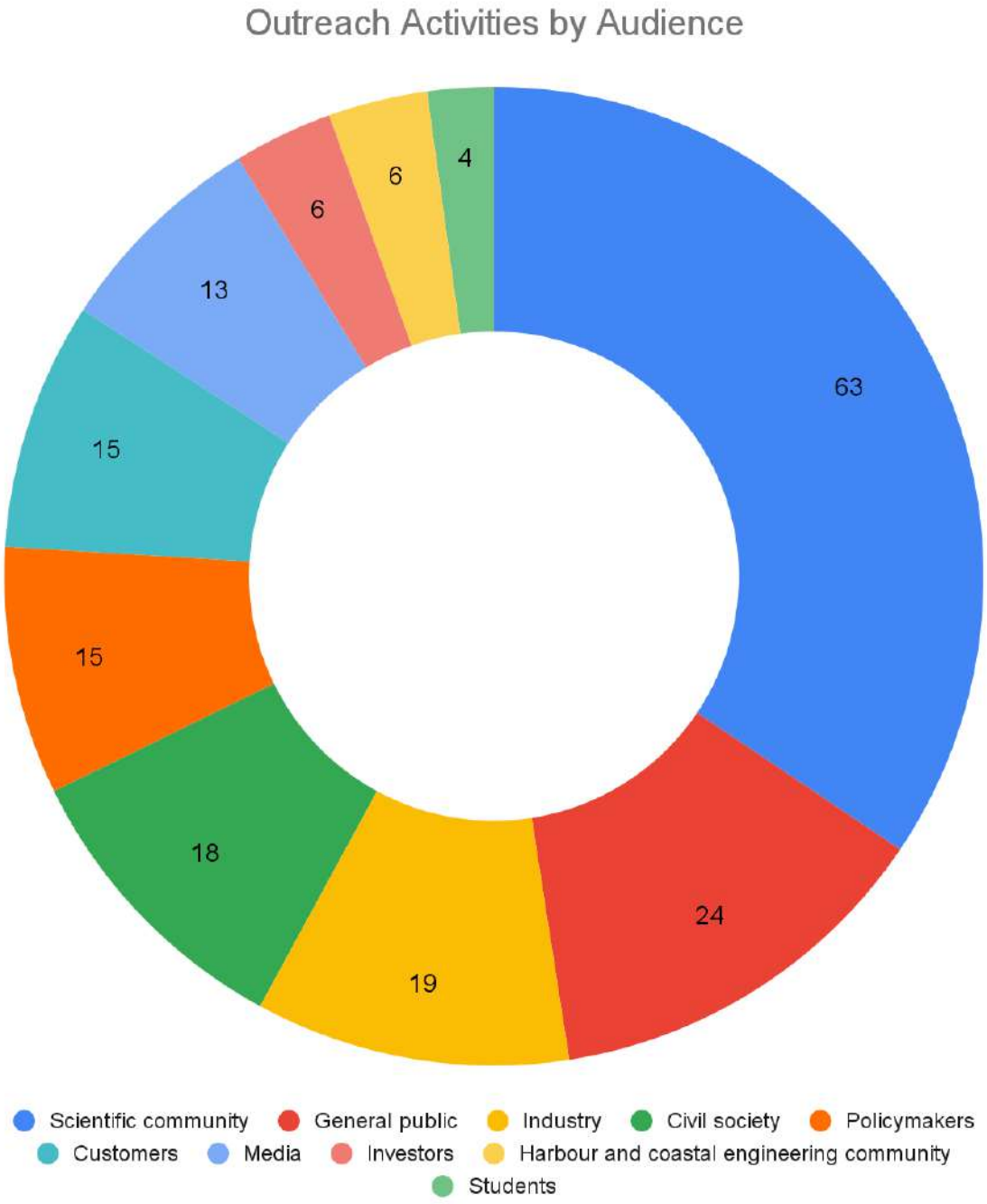
[58] https://aquadocs.org

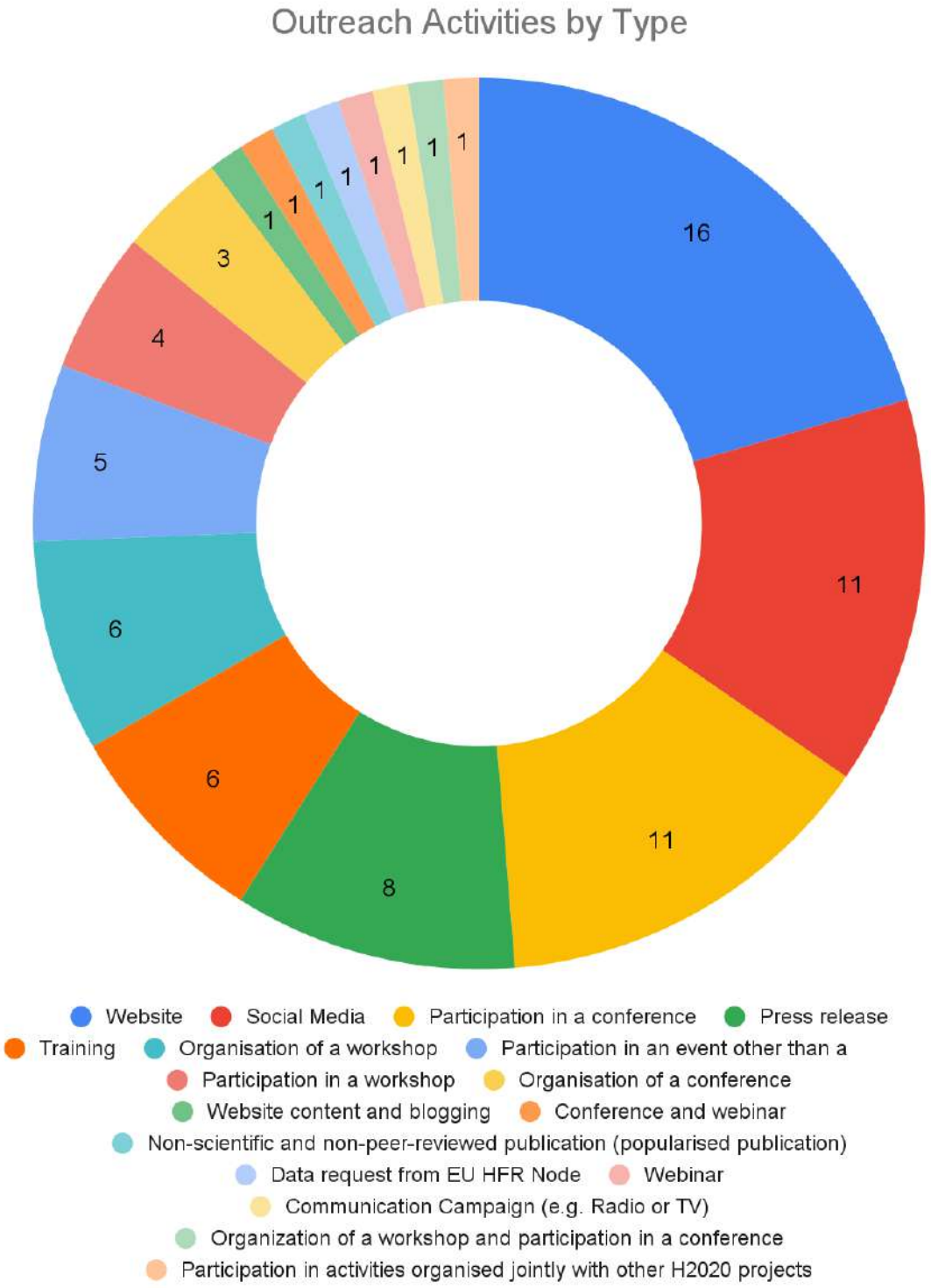
[59] https://classroom.oceanteacher.org

[60] https://repository.oceanbestpractices.org

[61] https://marinedataarchive.org


Outreach Activities by VA Infrastructure
10
8
7
6
5
5
4
4
4
4
4
4
3
3
3
3
2
2
1
1
PORTUS Observing and Forecasting System (ID 28.1)(PdE)
VLIZ Marine Data Archive (ID 38.1)(VLIZ)
CytoFluoTool (ID 8.1)(CNRS-LOG, ULCO)
OBPS-OTGA (ID 36.1)(IODE of UNESCO-IOC)
CefMAT (ID 6.1)(Cefas)
POSEIDON Multi Platform Observatory Data Center (ID 16.1)(HCMR)
SYKE-ALG@LINE (ID 34.1)(SYKE)
HIDROGRAFICO+ (ID 19.1)(IH)
SOCIB Data Centre Multi-Platform Observatory (ID 33.1)(SOCIB)
VOS Finnmaid GHG - BGC (ID 21.1)(IOW)
Utö Atmospheric and Marine Research Station (ID 15.1)
Swedish Oceanographic Data Centre (ID 32.1)(SMHI)
EU HFR Node/AZTI (ID 4.1)(AZTI)
Mawenzi (ID 1.2)(IFREMER ULCO-LISIC)
COSYNA (ID 17.1)(Hereon)
JERICO-ECOTAXA (ID 8.2)(CNRS-LOV)
JERICO RI e-Infrastructure (ID 33.2)(SOCIB)
Keri Island Research Station (ID 35.1)(TALTECH)
NorFerry/NorSOOP (ID 25.1)(NIVA)
1 more


Outreach Activities by Audience
63
24
19
18
15
15
13
6
6
4
Scientific community
General public
Industry
Civil society
Policymakers
Customers
Media
Investors
Harbour and coastal engineering community
Students



3.3. VA Concept

There is an opportunity to increase the access metrics of JERICO-S3 VA services by:

1. making it easier to find and access them

2. constantly monitoring access metrics to provide feedback to VA infrastructure managers of the impact of their actions.

Thus, the vision that inspires the VA-related activities of JERICO-S3 is to create a digital representation of the JERICO Research Infrastructure. This representation should facilitate both the enhanced (virtual) access to all JERICO resources, and the development of coastal research focused services.

In JERICO CORE, resources are to be understood as assets that contribute to the development of effective research and application activities which can be accessed by JERICO partners and by external users. They can be broadly classified into:

- Equipment
- Datasets
- Services
- Software
- Organisations, Projects, Observatories
- Support, Training, Manuals and best practices and similar assets
- Publications

This requires the construction of a central hub that becomes an easy and unique point of access to JERICO resources required to cover the whole ocean observing value chain (in line with EPOS, SERA and ACTRIS). This unified central hub should allow users to virtually find, access, manage and interact with JERICO resources. This hub should also be capable of recording and centralising access metrics to the VA services. This central hub is currently called JERICO CORE. Under WP7, Task 7.5 is developing a pilot version of JERICO CORE. Technical details of the implementation of JERICO CORE will be available in deliverable D7.6, due M34 (Jan 2023). JERICO CORE will be a VA service itself and it will fall under WP11's responsibility to operate and maintain it once its development finishes. It is expected to have the pilot version of JERICO CORE operational before the deliverable D11.3, due M40 (May 2023). The report on its implementation and operation will be prepared then.

The following is the presentation of the Virtual Access Concept, a framework for the long term strategy of the aforementioned central hub and its relationship with current VA services, and the rest of its stakeholders, in order to keep VA aligned with the long term goals of JERICO. This concept will be extended and further refined in WP3 of the JERICO-DS H2020 project.

3.3.1. VA before JERICO-S3

Before JERICO-S3, centralised access to JERICO data assets was achieved only via data aggregators (i.e: EMODnet Physics, CMEMS, SeaDataNet). Additionally, there was a

catalogue of VA services[62] on the JERICO website. VA services, which consisted of data access services exclusively, were monitored and received periodical feedback by a panel of experts (i.e: the work done under WP6 of JERICO-NEXT). However, there was no central point where a user could access all the coastal ocean information homogeneously. This means that the user was exposed to a large variety of technologies and was responsible for knowing where to find and how to access the information they were looking for.

3.3.2. JERICO CORE User Journey

Ideally, a person needing information about coastal oceanography should be able to search it in a central place. After an initial search, a set of results should be presented. Results could consist of the following:

- Equipment
- Datasets
- Services
- Software
- Organisations, Projects, Observatories
- Support, Training, Manuals and best practices and similar assets
- Publications

Upon selection of one of the results, more information on the selected resource should be displayed. These three pieces of information should be available:

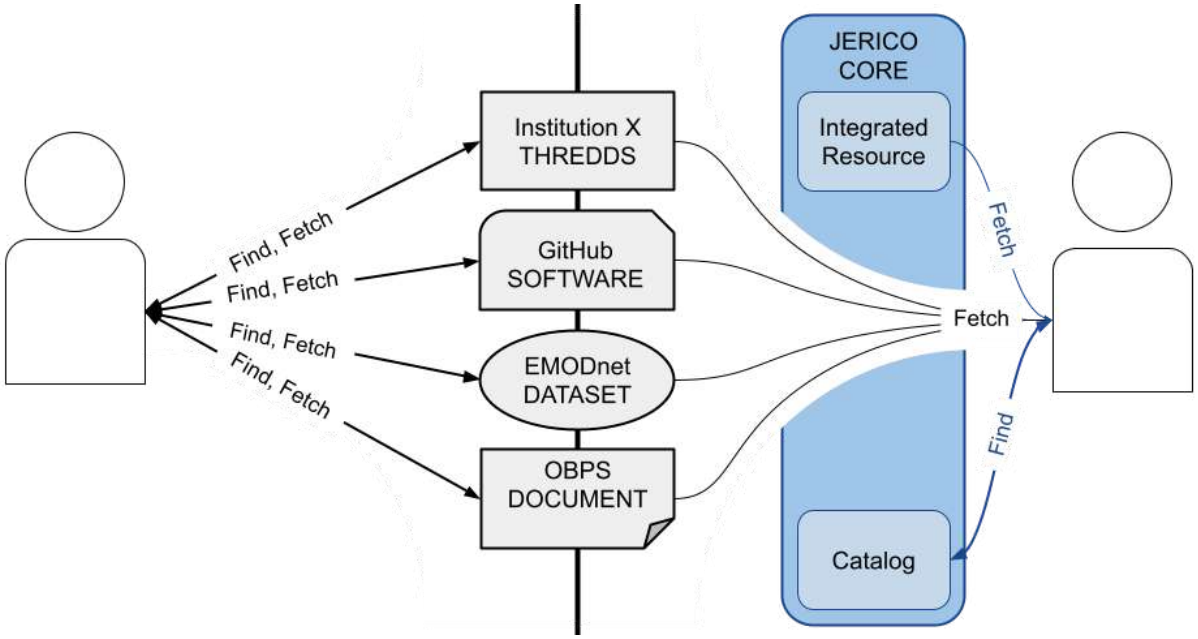
- Metadata of the resource: name, location (if applicable), relevant date range…
- A link to access the resource where it is hosted. Depending on the type of resource, this will be: a link to the original VA (or an aggregator) where the content is available, a link to an application form to access a physical platform, a link to an organisation's website…
- Links to related resources inside JERICO CORE. As an example, when viewing the information of a glider, this section should contain links to the following JERICO CORE pages: the organisation that owns the glider, the projects a glider has taken part in, the datasets produced with data from this glider, the best practices document used for the maintenance of the glider…

The key concept is that, once a resource has been found, it should be very easy to access it and discover related resources.

[62] https://www.jerico-ri.eu/virtual-access/

3.3.3. JERICO CORE Does Not Replace Original VA services

JERICO CORE will not replace existing VA services. JERICO CORE will not contain the data, only their metadata and a map of the relationships among resources. This means that, for users to virtually access a VA service they have found via JERICO CORE, they will have to access it where it is hosted (an institution server, a third-party hosting service like GitHub, an aggregator, a document repository…). In this regard, one can think of JERICO CORE as a Discovery Platform for coastal oceanography in the same way Google is a Discovery Platform for the web. Its role is to help the user find what they are looking for and send the user to the right place. As the following diagram shows, it is still possible to find and access the same VA services without JERICO CORE.



The above diagram represents two users who have to access the same hypothetical VA services. One of them uses JERICO CORE as a tool to find them, while the other needs to deal with the different underlying technologies to access them directly. Additionally, the user who chooses to use JERICO CORE has access to additional integrated services.

Users will still be able to access VA services directly, but that will put the burden of finding them on their own. On the other hand, a JERICO CORE user has a single central point of search that gives the necessary instructions to get the data from the original source.

3.3.4. Is This Yet Another Catalogue?

No. JERICO CORE feeds from a resources catalogue, but it is a metadata-rich, linked catalogue. This means that JERICO CORE understands the details of the resources (geographic distribution, valid date range, type...) and their relationships (which platform has produced which dataset, which best practices have been used to process which data…). The resources contained in the catalogue are very diverse and more resources will be

added. As an example, these are the types of resources that the catalogue could include: pieces of equipment, datasets, services, organisations, documents, people... JERICO CORE can use that information to help discoverability and to compose new products and services automatically.

JERICO CORE is a part of the research infrastructure that is not limited to any specific institution. Rather, it is connected to all JERICO institutions. This will allow the creation of integrated VA services that do not have a space to be managed in the current landscape. Specifically, it will allow the creation of services that integrate and present data collected and hosted by a variety of institutions.

JERICO CORE will automatically collect the metadata it needs for its catalogue from the available VA services. A set of standard vocabularies will be established for VA services to include so that JERICO CORE can learn about them and include them in the catalogue.

To be included in the Virtual Access Metrics System (VAMS), VA service managers will need to grant VAMS access to the metrics data source (Google Analytics, processed log files…).

3.4. Virtual Access Expert Panel

3.4.1. Description of VA Expert Panel

The access metrics that have been reported in this document will be reviewed according to Article 16.2b of the JERICO-S3 Grant Agreement. This article states that Virtual Access providers "*must have the virtual access services assessed periodically by a board composed of international experts in the field, at least half of whom must be independent from the beneficiaries*".

This deliverable (D11.1) consists of the collection of all the necessary information to evaluate all the Virtual Access services. The next deliverable (D11.2) will consist of the feedback provided by the aforementioned international experts. The third deliverable of WP11 (D11.3), due on M40 (May 2023) will collect, again, the relevant information to be reviewed again by the expert panel in M44 (Sept 2023).

3.4.2. Composition of VA Expert Panel

It has been the responsibility of the WP11 team to select and invite the members of the Virtual Access Expert Panel. It was decided to create a Panel of six members: three of them, independent from JERICO-S3 and three of them, members of JERICO-S3. The following is a short profile of each of the members of the Virtual Access Expert Panel.

- **Anca Hienola (FMI)**. Anca Hienola was awarded a PhD (Physics) in 2008 from the University of Helsinki, Finland. She works at the Finnish Meteorological Institute as a leading specialist and Open Science Taskforce leader. She has extensive experience in data flows and virtual data access metrics, including evaluation and metrics-based decision making and policies. She is involved in three major EOSC related projects - ENVRI-FAIR, EOSC Nordic and EOSC-Future.
- **Antonio Novellino (EMODnet Physics)**. Antonio Novellino has a PhD in Biotechnology and Bioengineering. He is the Research Manager at ETT Spa (Gruppo SCAI). He is a member of the EuroGOOS DATAMEQ group and contributes to several EuroGOOS Task Teams for advising on operational oceanography data management procedures and standards. He serves on the EMODnet Steering Committee, the EMODnet Technical Working Group. He is the EMODnet Physics coordinator and CMEMS DU deputy coordinator.
- **Thierry Carval (IFREMER)**. Thierry Carval is the head of Ifremer "Scientific Information Systems development". He is particularly involved in the management of in situ Marine data. He is the technical coordinator of the European service "Copernicus Marine in situ TAC" and data manager for the French JERICO sites. He is involved in the development of EOSC through projects such as ENVRI-FAIR, EOSC-HUB, EOSC-FUTURE or BLUE-CLOUD. He ensures that the JERICO VA services are well aligned with the EOSC E-infrastructure.
- **Simon Keeble (Blue Lobster)**. Simon Keeble is the Managing Director at SME Blue Lobster IT Limited based in Wales, UK. He has extensive experience in big-data systems and web / mobile interfaces and user interaction within the marine and environmental sciences sectors. The Blue Lobster team operated one of the first services to be part of Virtual Access, in the Framework 7 project FixO3, and they continue to develop and operate a variety of web-based platforms presenting a variety of observational, model and complex derived product data.
- **Alan Deidun (University of Malta)**. Prof. Alan Deidun is a marine biologist by training. He is Malta's first-ever Ocean Ambassador and, as of September 2019, he is a member of the EU Commission's Ocean Mission Board. He directs the Malta Training Centre of the International Ocean Institute (IOI) and he is a Fellow of the Royal Society of Biology (London). He is also a Full Professor within the Faculty of Science of the University of Malta and he is currently a contributor to the UN's Second World Ocean Assessment.
- **Sebastien Mancini (AODN-IMOS)**. Since August 2008, Sebastien Mancini has worked as a project officer, data services team leader and, more recently, as the Director of the Australian Ocean Data Network (AODN), which is in charge of the ingestion, curation and preservation of all data and relevant metadata collected by IMOS. During this period, he has contributed to the creation and improvement of the single integrative framework for data and information management that allow discovery and access of the data by scientists, managers and the public.

4. OUTREACH, DISSEMINATION AND COMMUNICATION ACTIVITIES

As of the publication date of this document, VAMS is not still publicly accessible. Efforts to integrate all VA providers included in WP11 into VAMS to produce the most complete deliverable have been prioritised.

VA providers and members of the VA Expert Panel have, however, access to their dashboard, as it is a great tool to monitor the evolution of each infrastructure.

VAMS will probably be publicly accessible in the future. Discussions about the details of the publications are yet to take place.

5. CONCLUSIONS

A Virtual Access Metrics System has been successfully developed, which allows us to analyse, in real-time and in a centralised place, access metrics from the different infrastructures. This has had many positive effects:

- Over the long term, preserving log files from servers has proved to be challenging. Different VA providers have experienced similar incidents where, due to changes in the backup policies or directories or due to a change in the output directories of logs, these files have been inadvertently lost. Having a system that constantly monitors the files and analyzes them, reduces the dependency on these files.
- For some VA providers, there was no system in place to get analytics on the access to their infrastructure. For them, VAMS provides the first opportunity to monitor their systems, detect issues and plan upgrades.
- Having a single point of access to the metrics provides an efficient way to evaluate the VA infrastructures that JERICO members provide

This exercise has, also, shown some limitations that could be overcome to provide even a better system:

- It is hard to assess the data obtained via Google Analytics, as a standard configuration for Google Analytics views was not agreed upon before the beginning of the project. This configuration cannot be added *a posteriori*, as they filter whether pageviews are saved or discarded. This configuration allows us to discard traffic from robots and from internal users. We should ensure that we are using the automatic filter for robots. Also, a common policy of how to handle the measurement of internal traffic would be beneficial for the system.
- Some of the infrastructures' data can be obtained via aggregators (i.e: CMEMS, EMODnet, SeaDataNet…). Being able to measure access via aggregators would be a step further in the direction of changing the focus from measuring access to VA services (a specific THREDDS data server, a specific website…) to measuring access to the content generated by an institution regardless of where it is hosted. JERICO CORE will help build that capability. Standardising metrics so that one can sum metrics of different sources would be another good step.
- The usage of software has proven particularly difficult to measure. While measuring the downloads of software is feasible, measuring the usage would be more interesting. In the future, potentially, software that generates output files with metadata will be measurable by checking how many of those files are found by JERICO CORE.

Regarding the outreach activities, having the responsibility and the tool to report outreach actions has increased the awareness of the need to perform them among VA providers. So, more activity is expected in the future.

6. ANNEXES AND REFERENCES

Annexe 1: VAMS Dashboards from each Partner

Next, the dashboards of each VA service of JERICO-S3 WP11 are presented.

Annexe 1.1 Mawenzi (ID 1.2)

Access to Mawenzi software packages is monitored in two different ways: the CRAN downloads API and the log files of the documentation website. The following presents the two alternatives.

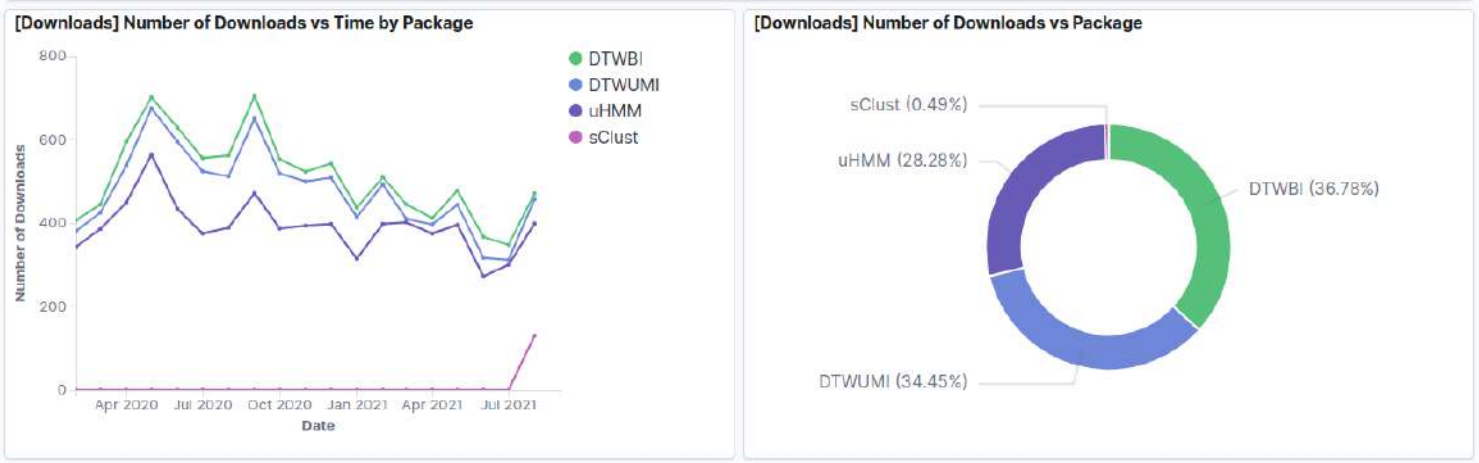
Downloads from CRAN

The data underlying the following plots have been obtained from the CRAN downloads API.

First, the overall numbers of downloads will be presented. Next, the temporal evolution of downloads and share of each package is analysed.

In the following image, the total number of downloads of each package is shown.

In the next image, the number of downloads of each package over time and the proportion of downloads of each package overall are shown.



There is a remarkable correlation between the number of downloads of DTWBI, uHMM and DRWUMI, which suggests that these packages are being used together.

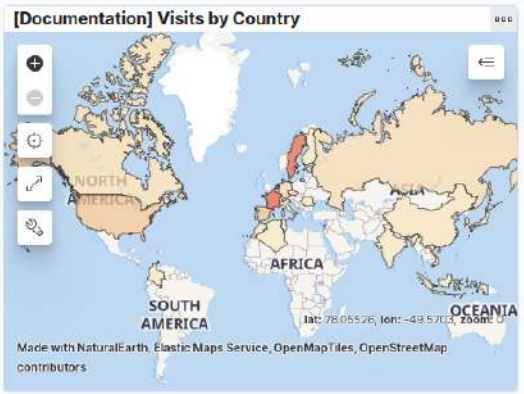
Visits to the Documentation Page

Software usage cannot only be measured by downloads, as a piece of software can be downloaded once and used many times. Therefore, It has been decided to measure access to the documentation pages too, to be able to estimate the trend in the usage of the software.

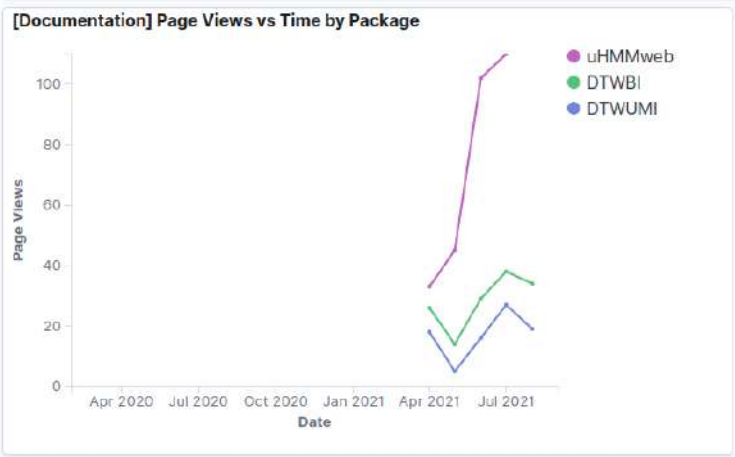
The data underlying the following plots have been obtained from the log files of the server of mawenzi's website. Unfortunately, logs of the first months of JERICO-S3 had been lost. Only data from April 13, 2021 is available.

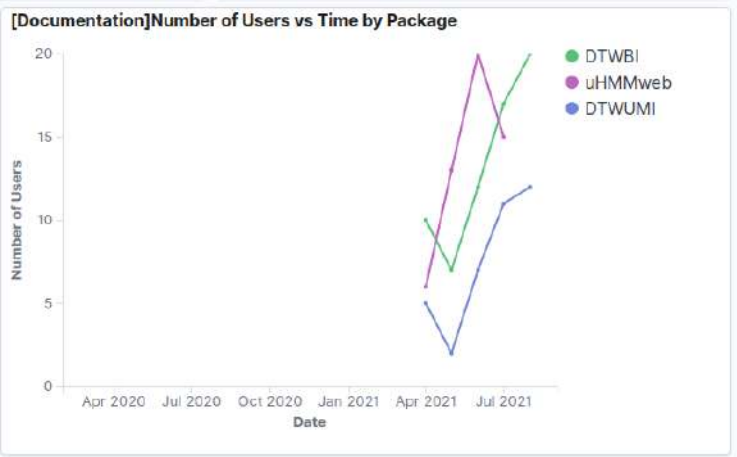
First, the general metrics and geographic analysis of the visits are presented. Then, the temporal evolution of the traffic is shown. Finally, the content visited and the origin of traffic is analysed.

In the next image, the overall number of visits, users and countries who accessed each package's documentation and the geographic distribution of visits are presented.

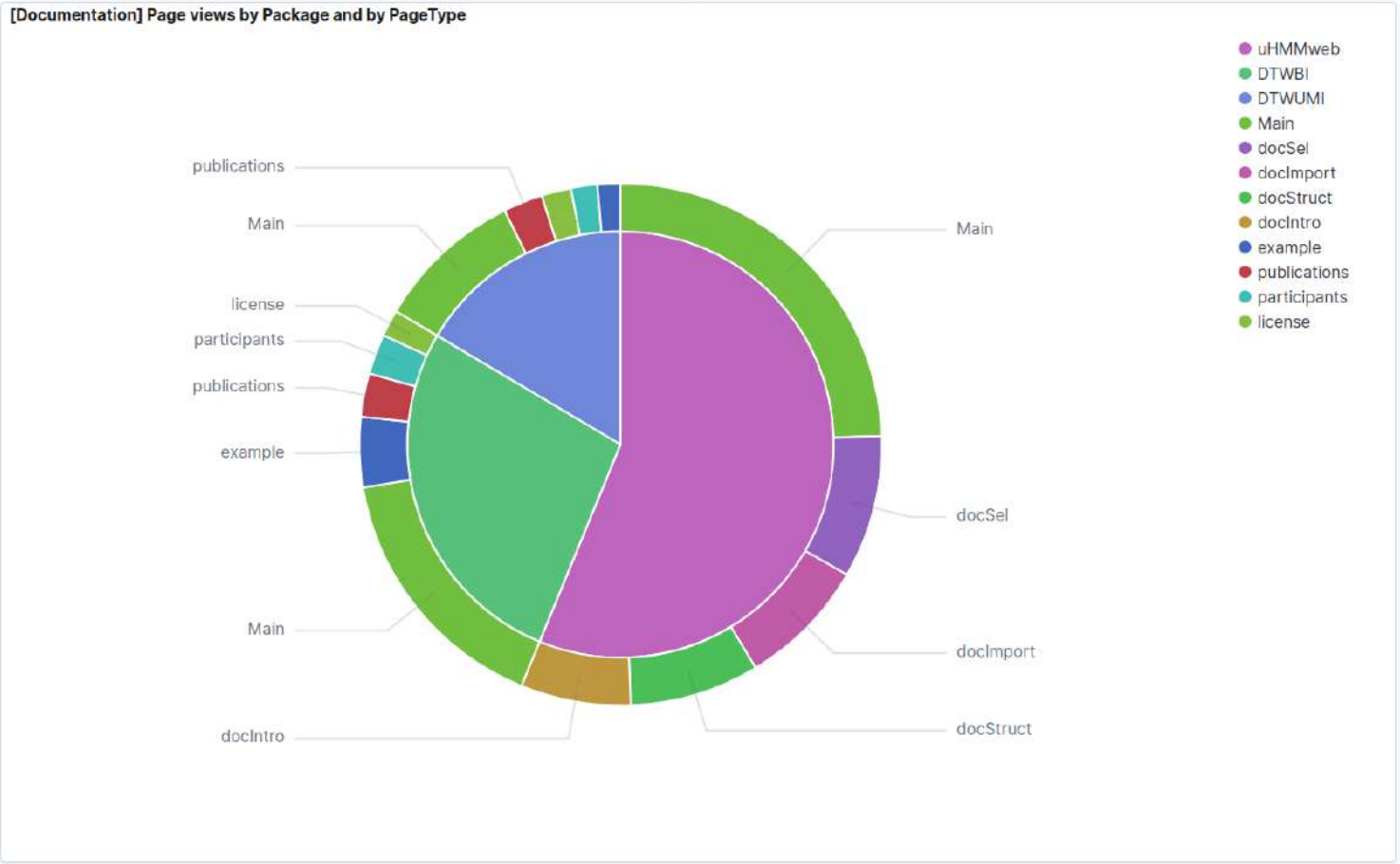


Next, two plots that show the temporal trend of pageviews and users respectively for each package are shown.

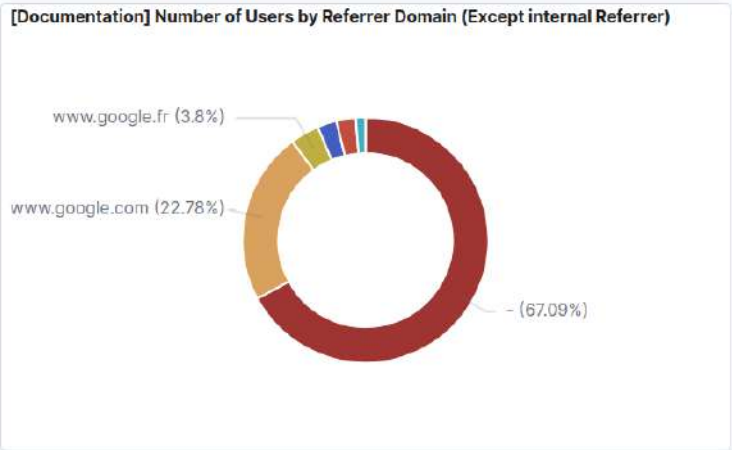


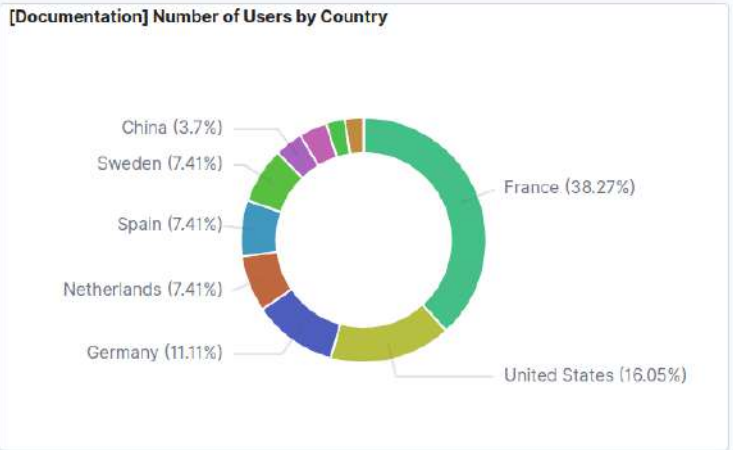


The following image shows a plot with the number of pageviews that each page of each package received.



Finally, the traffic origin is analysed: by both referrer ("-" means referrer is not informed, like a visit from a bookmarked link or direct access via software) and country.





The same correlation between the uHMMweb and DTWUMI that was seen in the previous dashboards (analysis of the number of downloads of the software) can be observed. Also, the number of visitors that do not have a referrer domain is remarkable. This is interpreted as users who have bookmarked the links. Finally, the geographic distribution shows interest

from Europe.

Annexe 1.2 EU HFR Node/AZTI (ID 4.1)

The data underlying the following plot have been obtained from Github API. Data is only available from June 18th 2021 because Github only keeps access metrics for 15 days. Therefore, data regarding access 15 days before the development of VAMS was not kept.

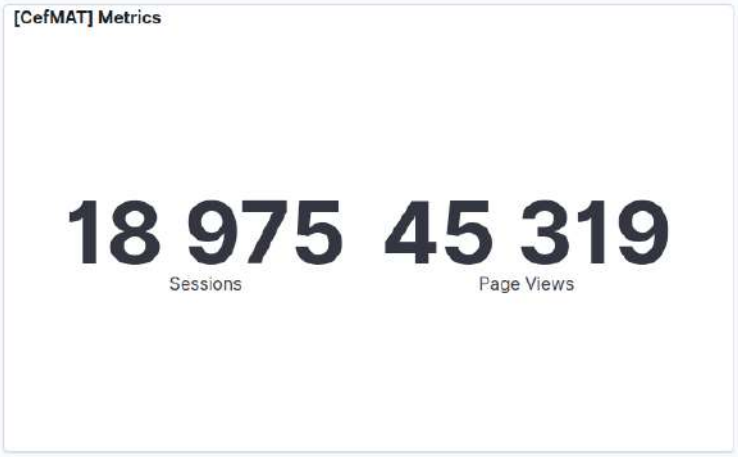
The following plots show the number of Views and Unique Views the software package JRadar has had on Github.

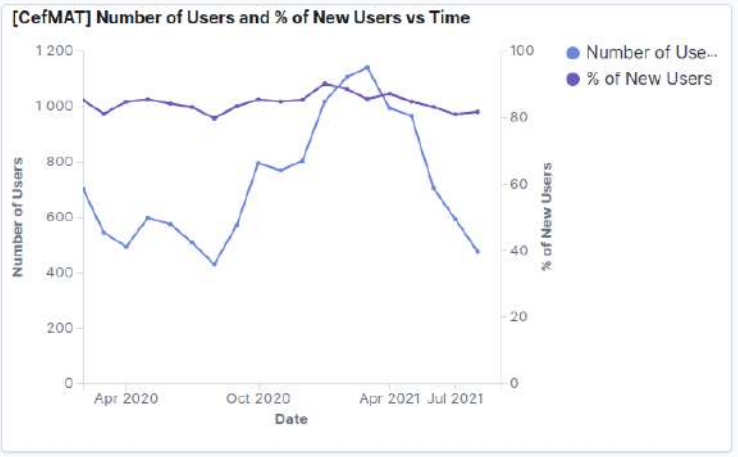
Annexe 1.3 CefMAT (ID 6.1)

The data underlying the following plots are obtained via a Google Form that CefMAT staff kindly fill monthly based on their Google Analytics reports.

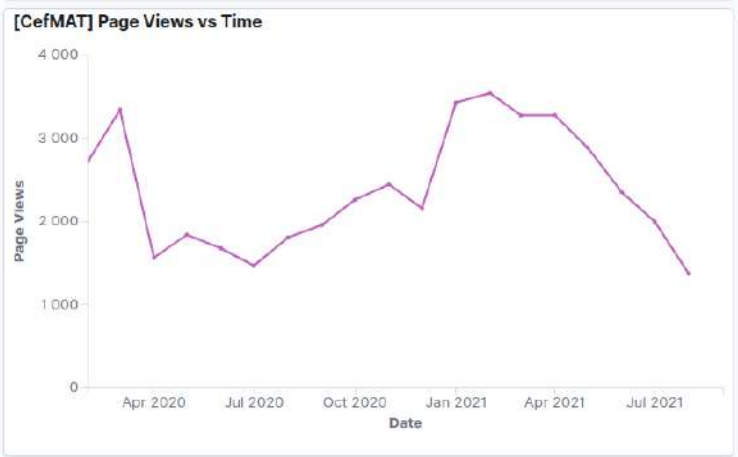
First, the volume of traffic will be analysed. Then, metrics related to session quality (session duration and bounce rate) will be shown. Finally, the origin of traffic will be analysed.

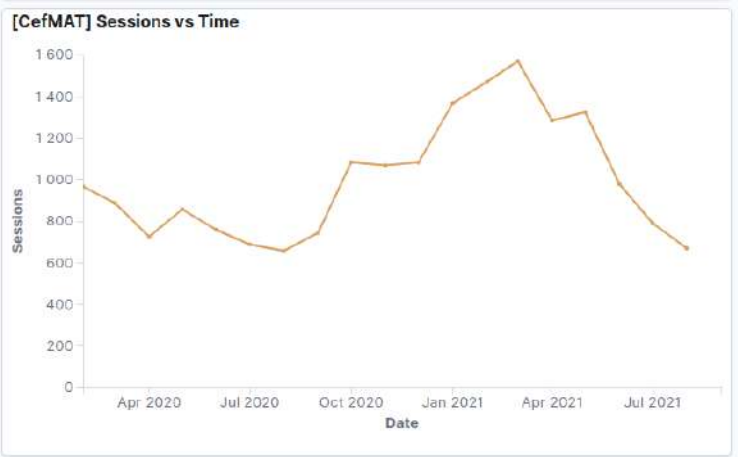
In the next image, a numeric representation of the access metrics and the users analysis is presented.



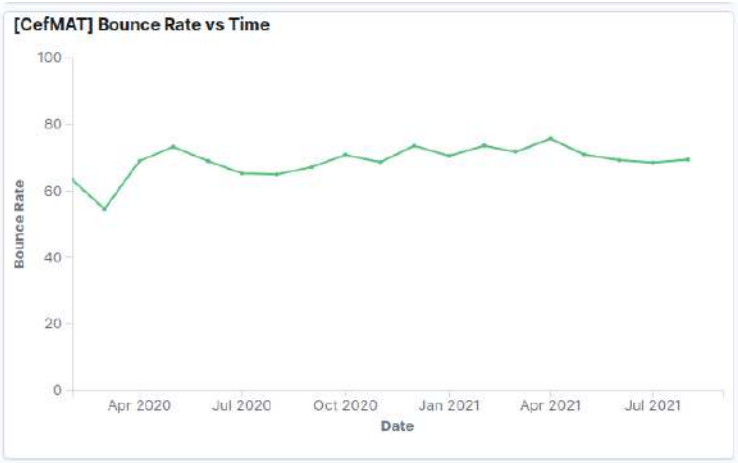


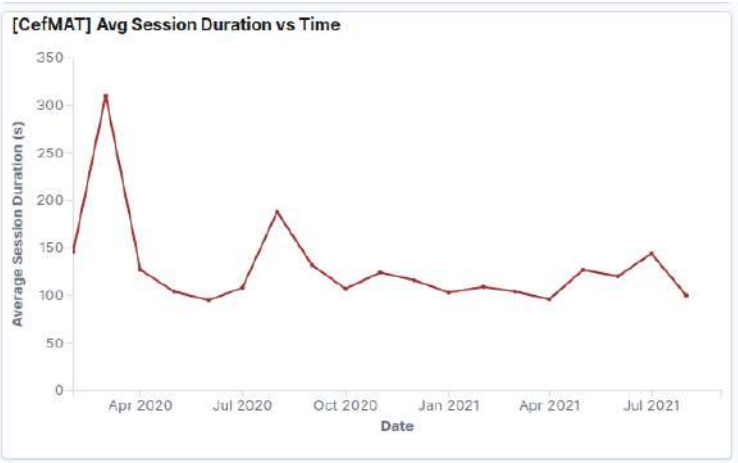
Next, a temporal analysis of the pageviews and sessions has been added.



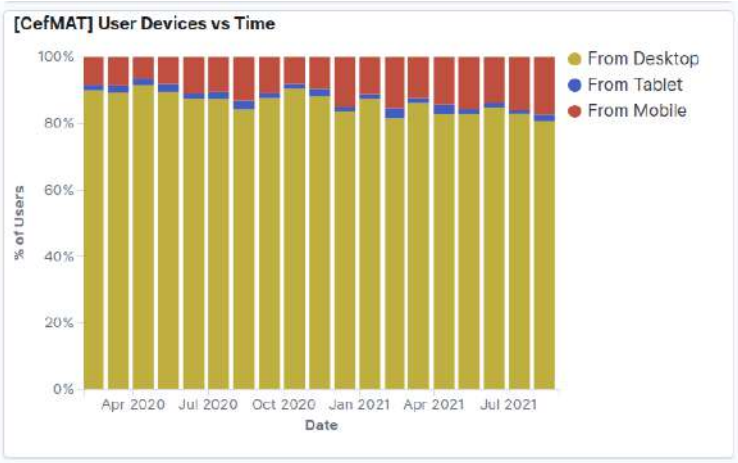


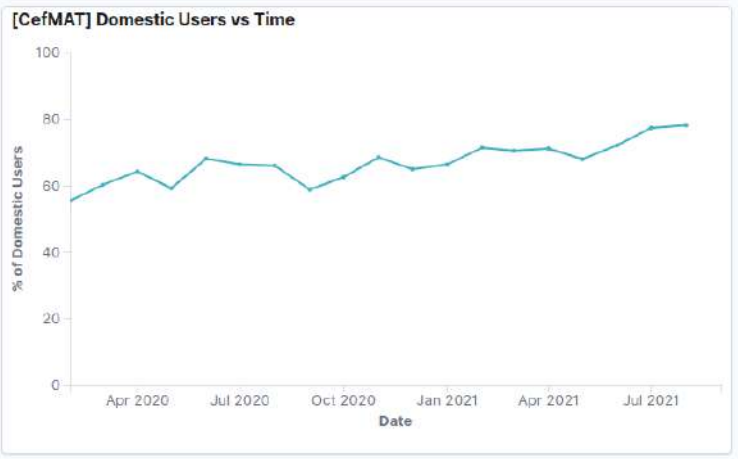
In the following image, metrics related to the on-page user experience: bounce rate and session duration are analysed.





Finally, the user origin is analysed: both by device and country of origin.





As expected by the VA infrastructure managers, some seasonality in the usage of the website is observed. Around spring, the number of users, sessions and pageviews increases. The bounce rate and average session length, however, do not seem to change, suggesting that the way the site is used is not seasonal.

Annexe 1.4 CNR TirLig e-infrastructure (ID 7.1)

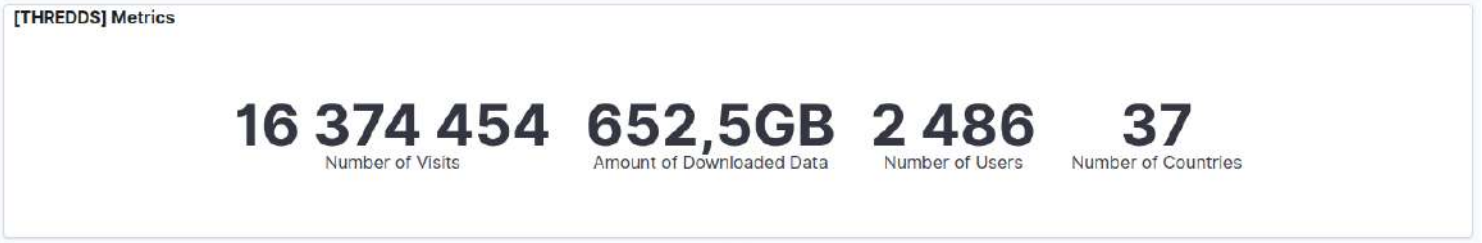
This VA infrastructure consists of two parts: a THREDDS data server that we monitor by analysing its log files and a website monitored via Google Analytics. Next, the access metrics of each of these parts are presented.

THREDDS Data Server

The data underlying the following plots have been obtained from the log files of the THREDDS data server.

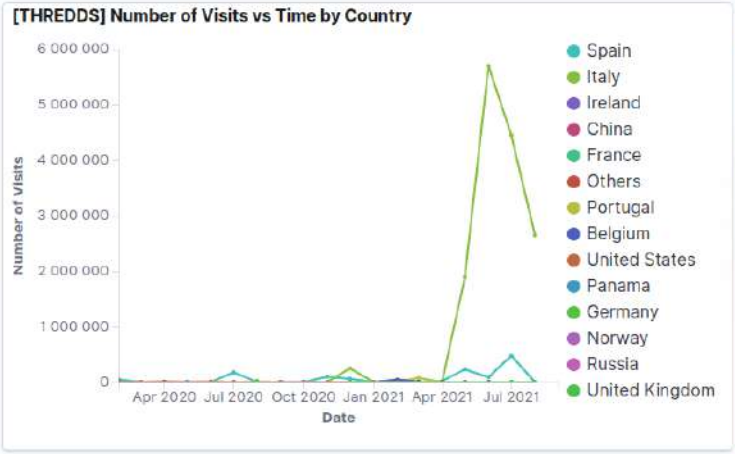
First, the overall access metrics will be presented. Then, traffic by THREDDS protocol, geographic origin, used client, referrer domain and operating system will be analysed.

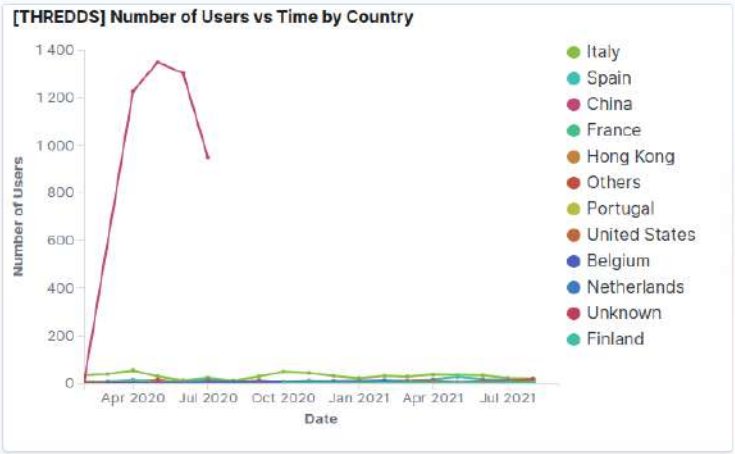
In the next image, the general access metrics in a numeric form are presented.

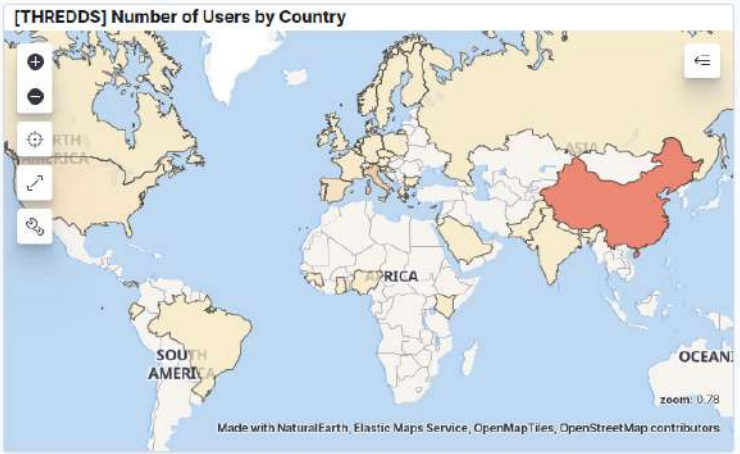


The following image contains the analysis of the traffic by THREDDS protocol (traffic to the catalog HTML pages is assigned to protocol "Catalog" and traffic to the HTML pages of viewers is assigned to "Viewer" protocol).

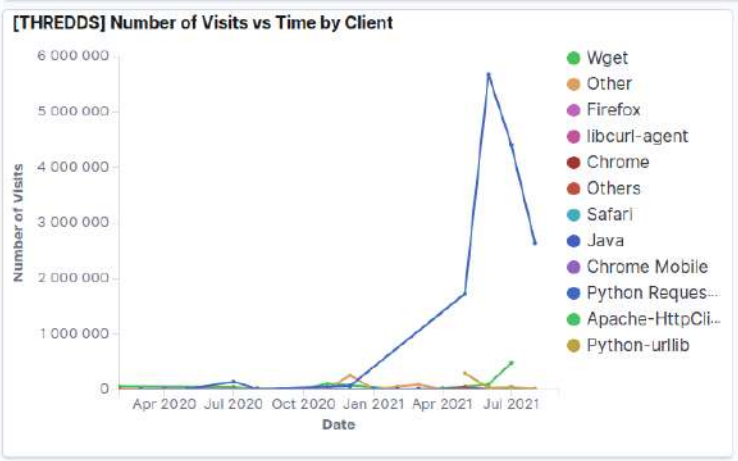
Next, plots and maps with a geographic analysis of the traffic seen by the server are shown.

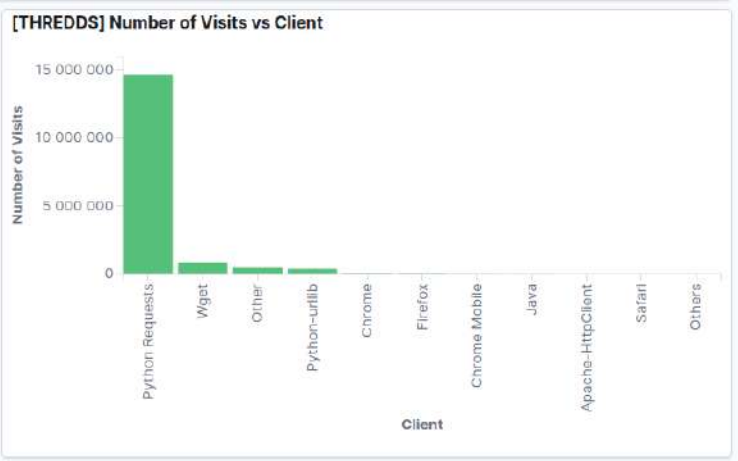


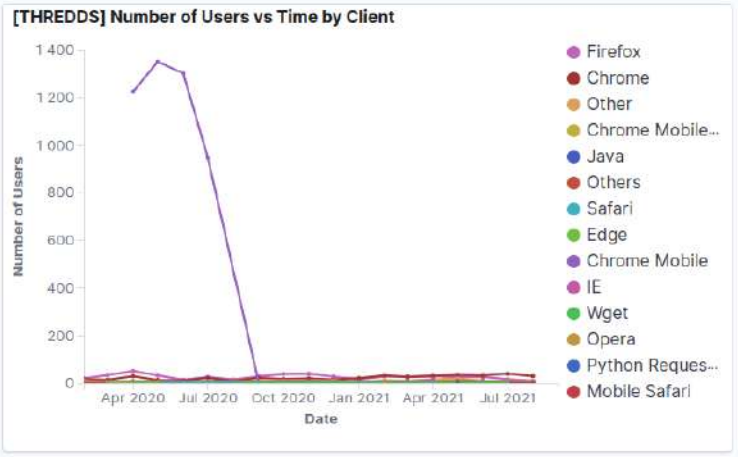


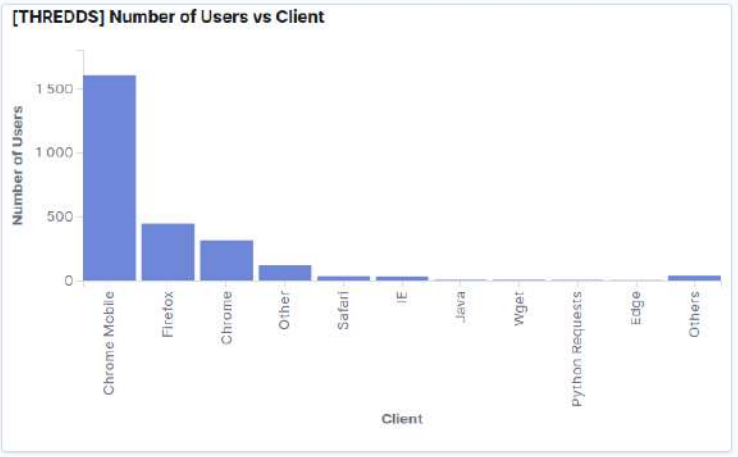


The fourth image shows an analysis of the traffic based on the software used by the user to perform the request (a browser or an HTTP library).

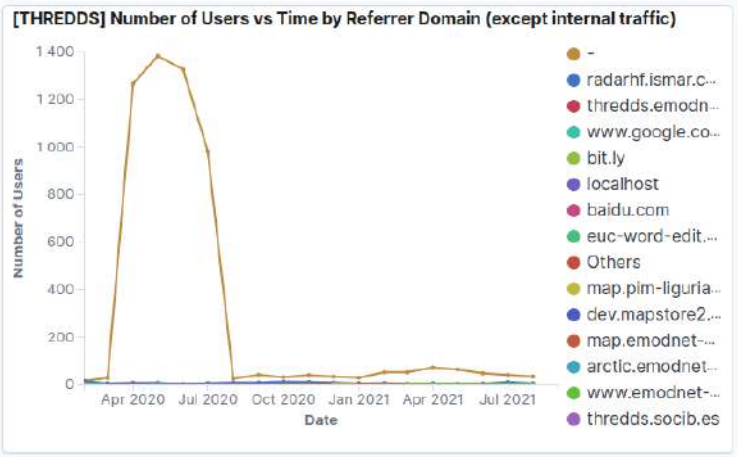


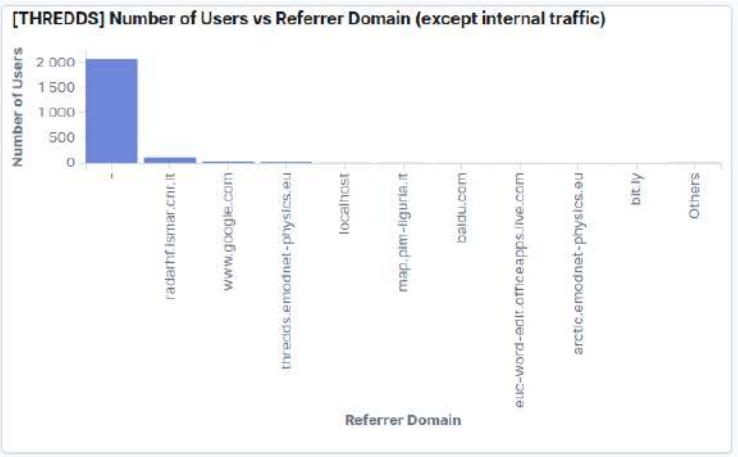






The image below contains an analysis based on the referrer domain ("-" means referrer is not informed, like a visit from a bookmarked link or direct access via software).

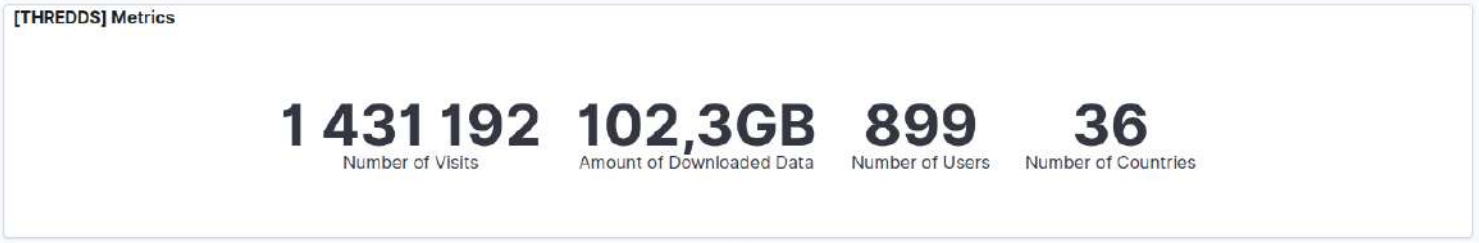




Finally, an analysis of traffic based on the operating system of the user is presented (that information is extracted from the browser version, for HTTP libraries, the information is not available and it is contained in the "Other" category).

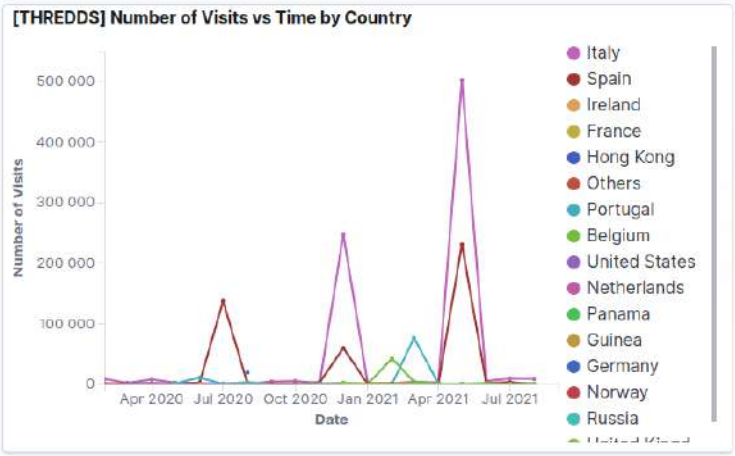
It can be observed that many of the plots are dominated by big, short peaks. One of the peaks is in the number of users around the summer of 2020. There are two other peaks in the amount of data in spring 2021. In order to see the trend of the baseline activity, the same images are included next where traffic from China (responsible for the users' peak) and two IPs that are responsible for the peaks of the amount of data (one is Italian, the other one is Spanish and belongs to AZTI) have been filtered out. The traffic from China appears to be a robot that scrapped the catalogue of the THREDDS data server using many different IPs (hence, creating a peak in users). The traffic from the excluded IPs seems to be legitimate scientific downloads.

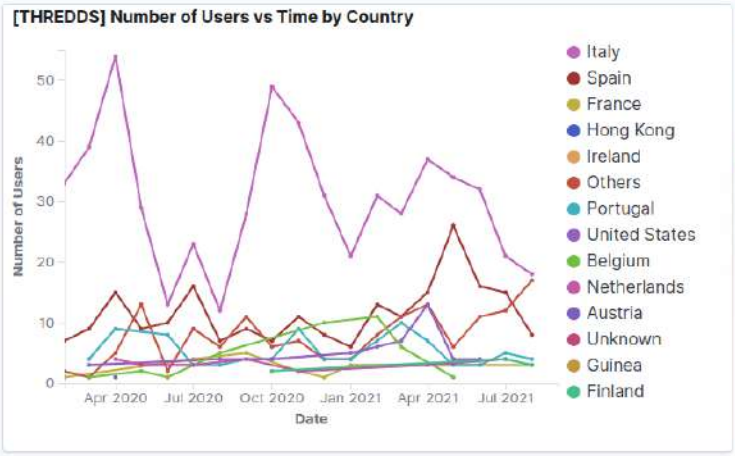
In the next image, the general access metrics in a numeric form are presented.

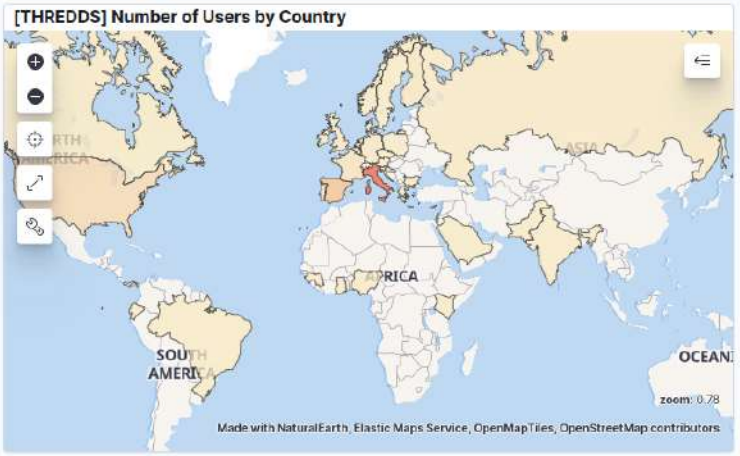


The following image contains the analysis of the traffic by THREDDS protocol (traffic to the catalog HTML pages is assigned to protocol "Catalog" and traffic to the HTML pages of viewers is assigned to "Viewer" protocol).

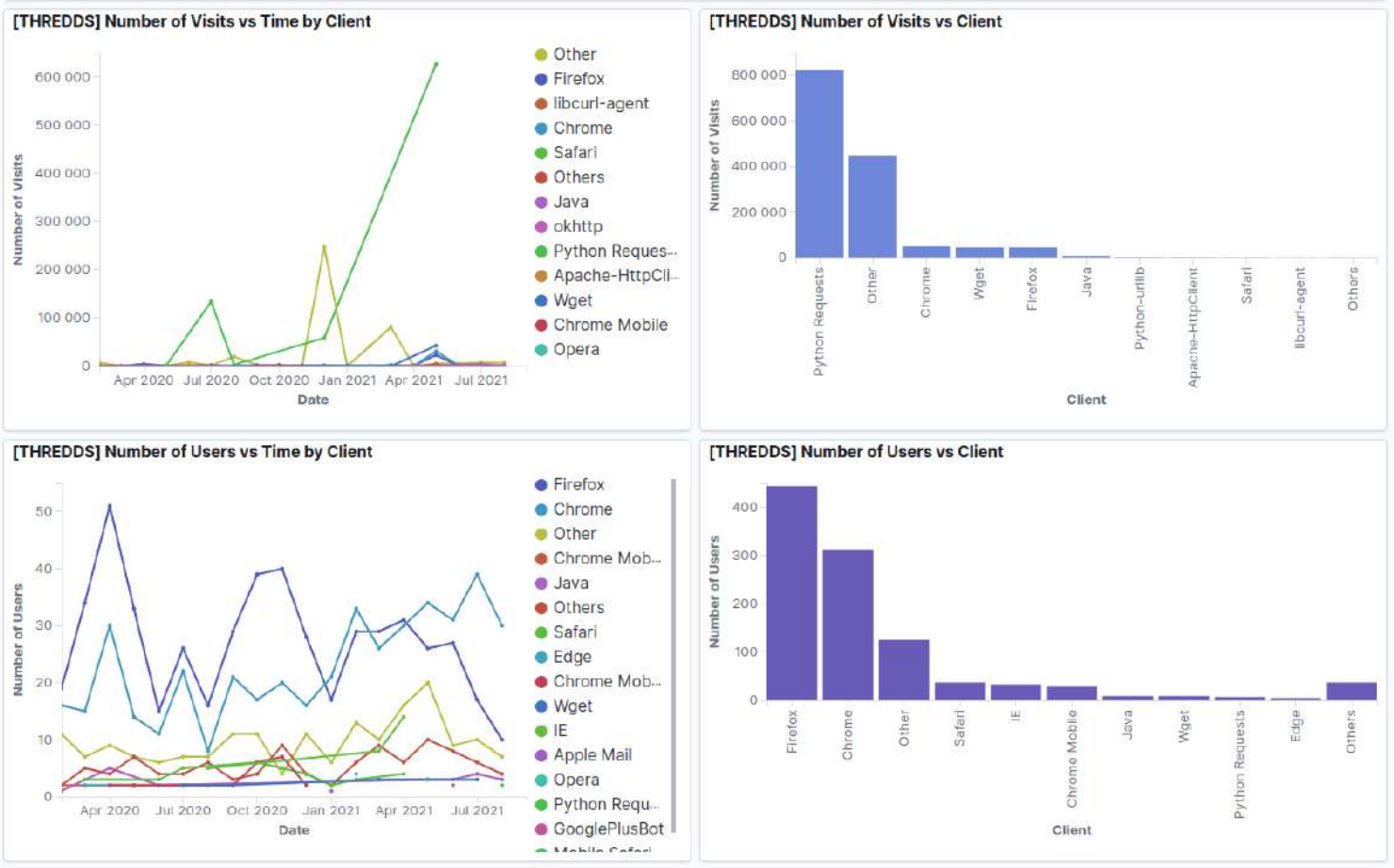
Next, plots and maps with a geographic analysis of the traffic seen by the server are shown.



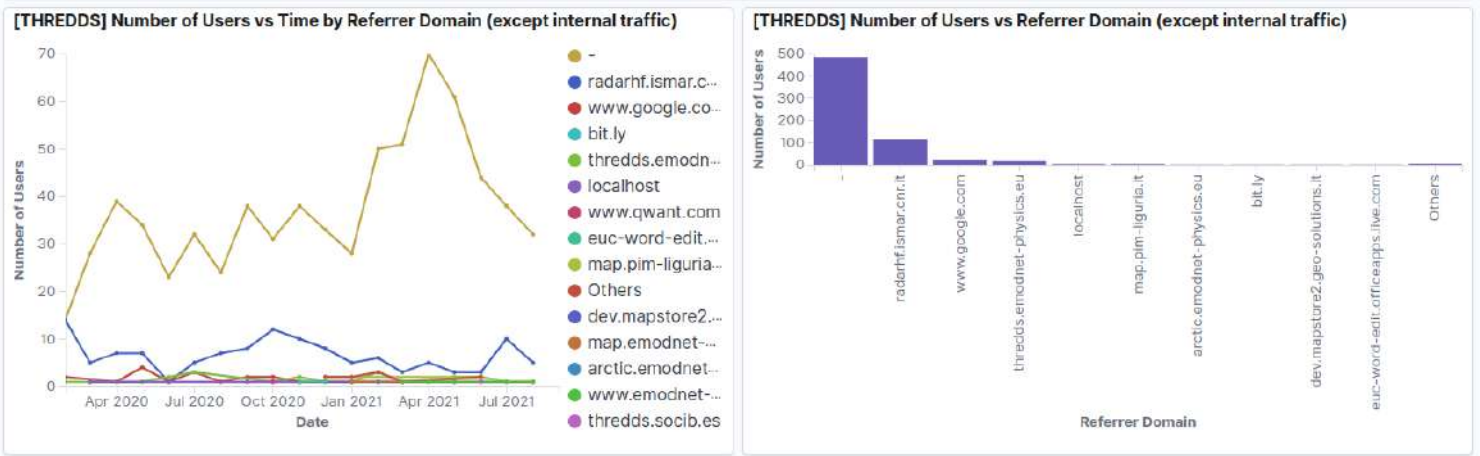




The fourth image shows an analysis of the traffic based on the software used by the user to perform the request (a browser or an HTTP library).



The image below contains an analysis based on the referrer domain ("-" means referrer is not informed, like a visit from a bookmarked link or direct access via software).



Finally, an analysis of traffic based on the operating system of the user is presented (that information is extracted from the browser version, for HTTP libraries, the information is not available and it is contained in the "Other" category).

In general, it can be concluded that the baseline activity has no growing or decreasing trend.

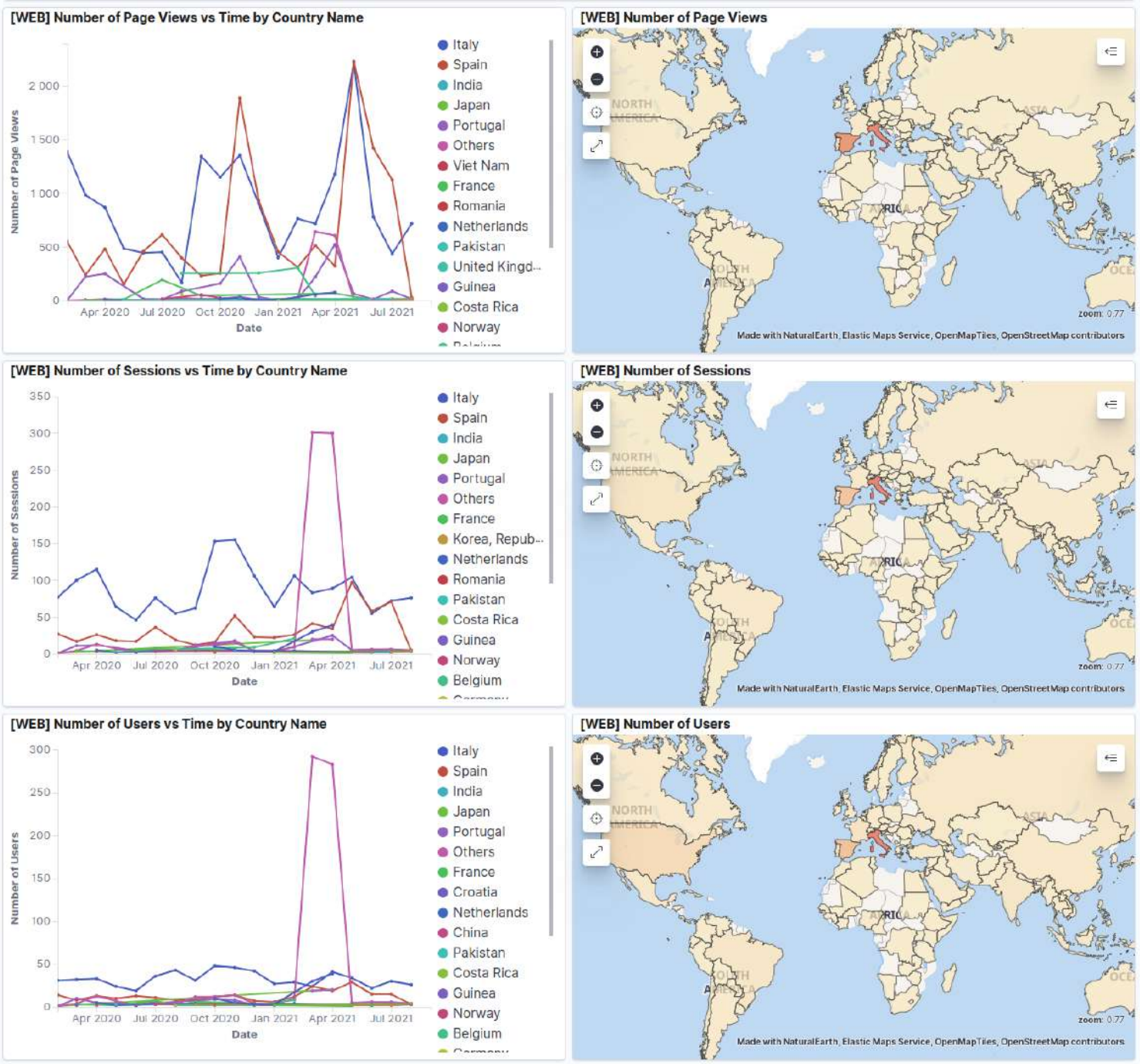
Website

The data underlying the following plots have been obtained from the Google Analytics API.

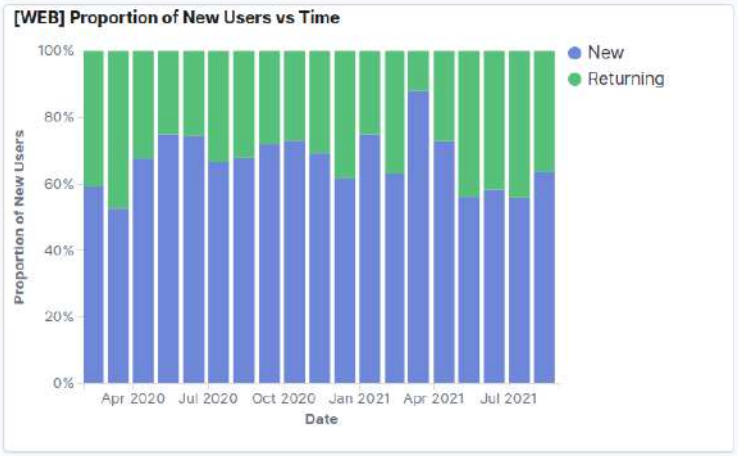
First, the overall access metrics will be presented. Then, the following analysis are included: geographic analysis, analysis by user type (new vs returning), bounce rate analysis, user device analysis and traffic source analysis.

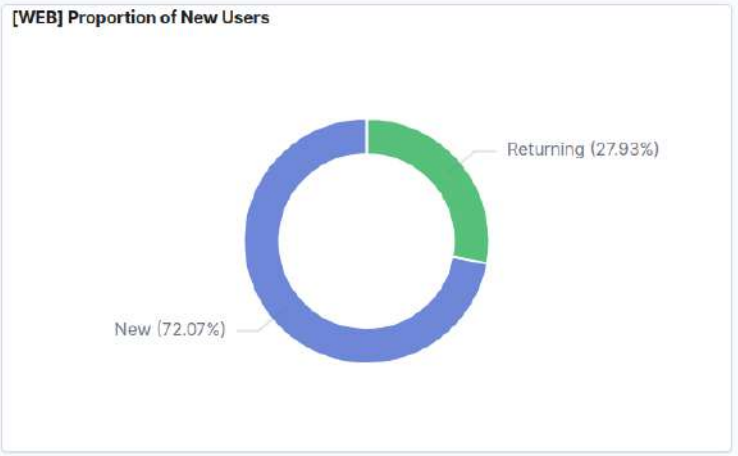
In the next image, the general access metrics in a numeric form are presented.

The following image contains the geographic analysis of the traffic.

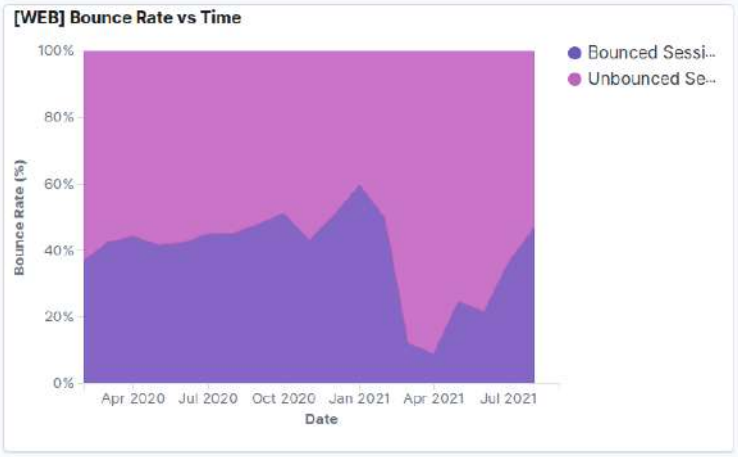


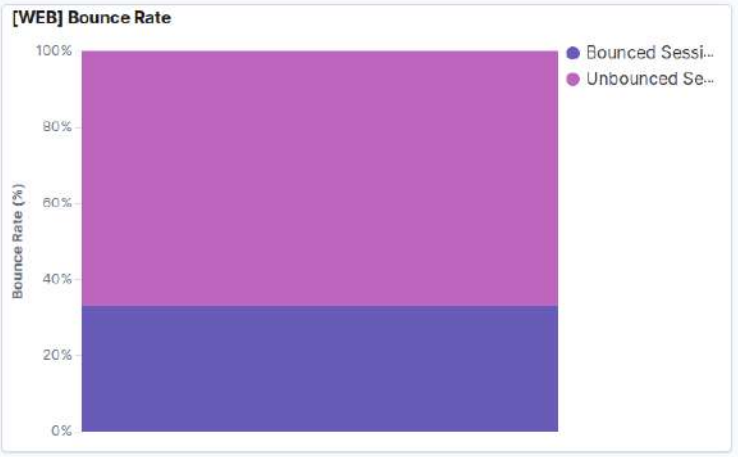
Next, plots analysing the proportion of new users are shown.





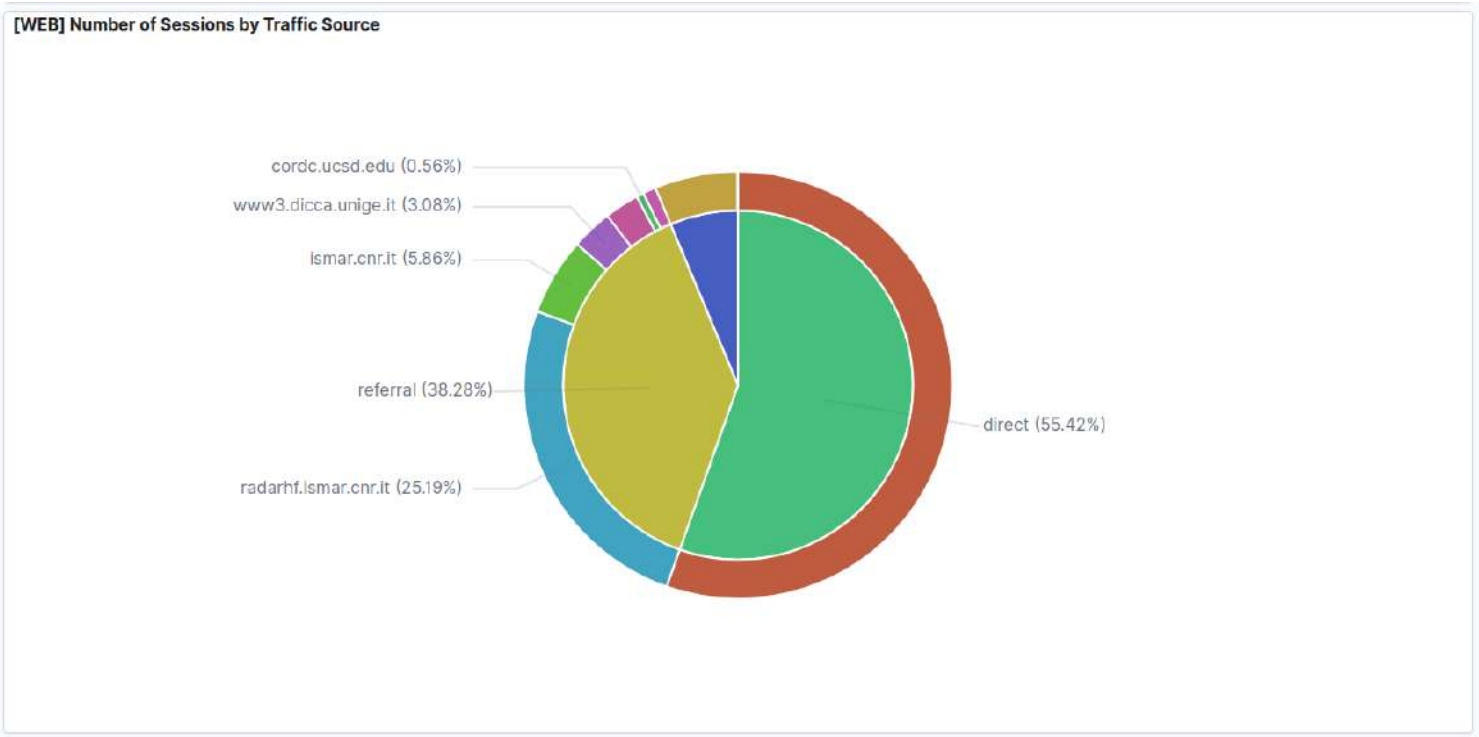
Next, the bounce rate is analysed.





The image below shows the proportion of sessions that come from each type of device and the influence of the type of device on the bounce rate.

Finally, an analysis of traffic based on the source is presented.



Annexe 1.5 EU HFR Node/CNR (ID 7.2)

This VA infrastructure consists of two parts: an HF Radar file processing service and an HF Radar processing software. Next, the access metrics of each of these parts are presented.

Processing

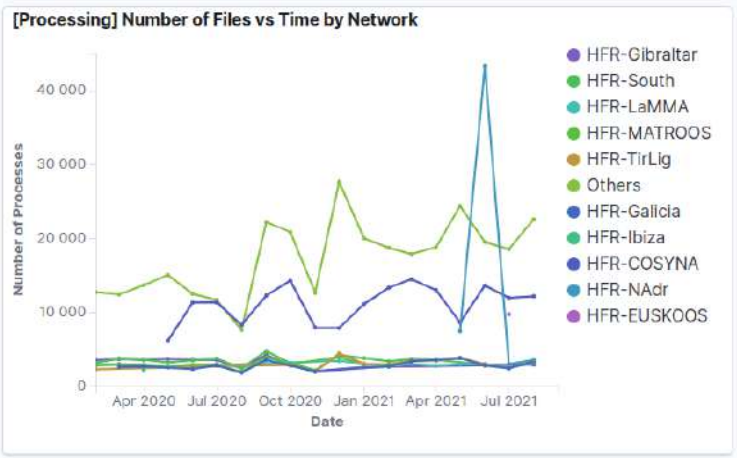
The data underlying the following plots have been obtained from a database that contains the registry of each time a file is sent to the processing service and each time a file is output by the processing service. There are three cases:

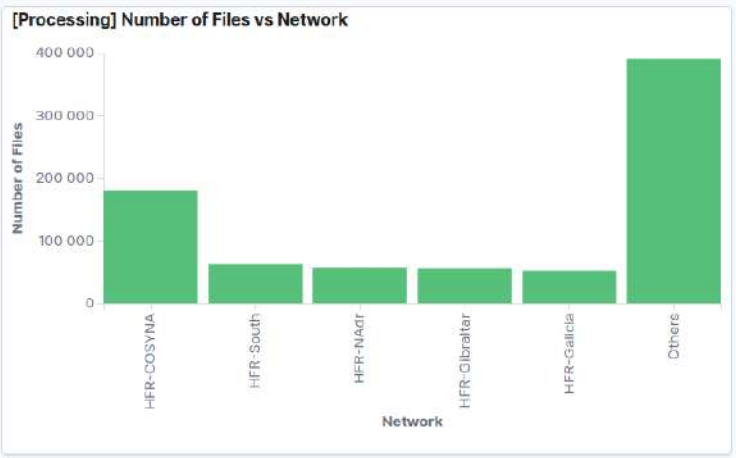
- Both the input and output registry of a process are present. This is the most common case and it allows us to calculate the amount of time a process took.
- Only the output registry is present: This happens when the processing service takes input files from a THREDDS data server. It is normal, but it does not allow us to calculate the amount of time a process took.
- Only the input registry is present: this means the process has not finished. It can be because it is still running or because the process crashed. This case does not allow us to calculate the time a process took.

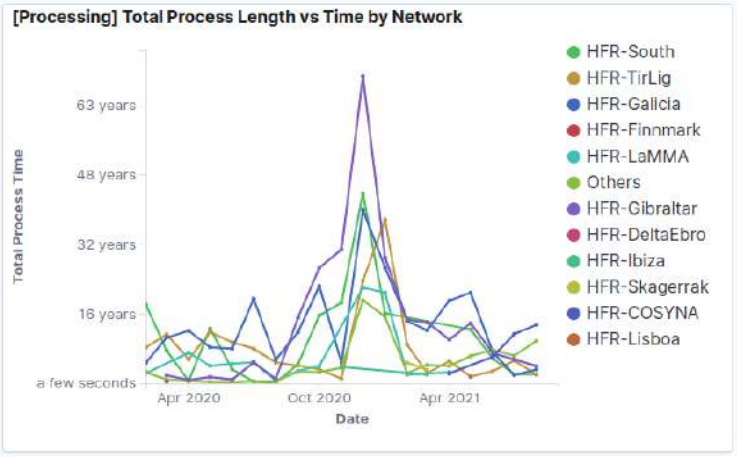
First, overall metrics will be presented. Later, the following will be presented: analysis by HFR network, analysis by process type and status and analysis by file size.

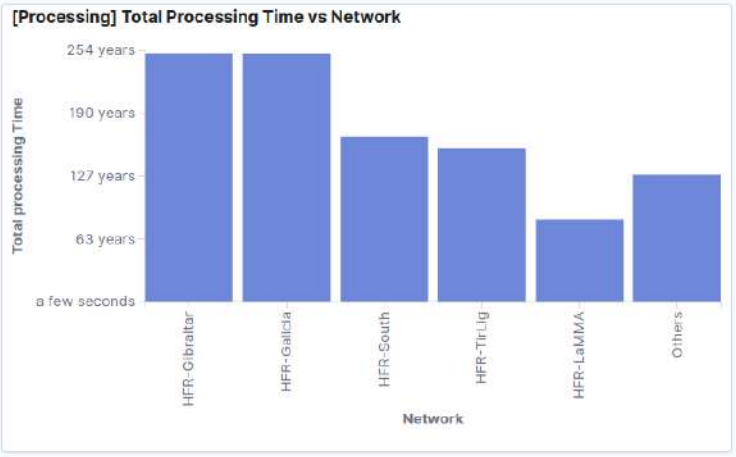
In the next image, the general access metrics in a numeric form are presented.

The following image contains the analysis of the processes by the network of HF Radar that generated the data

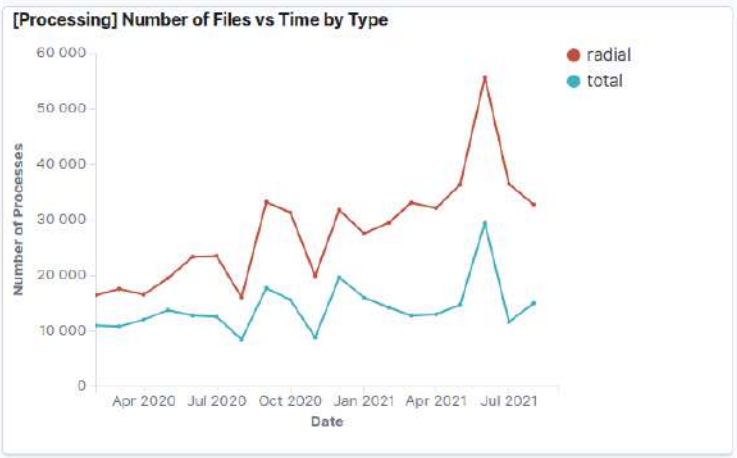


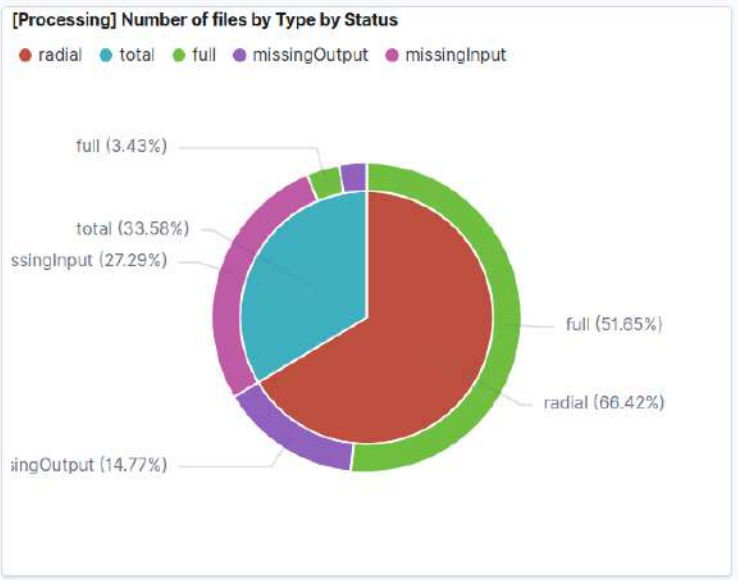


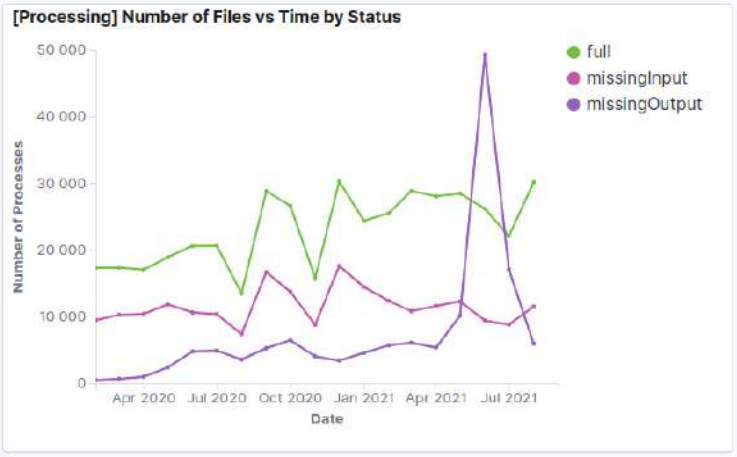


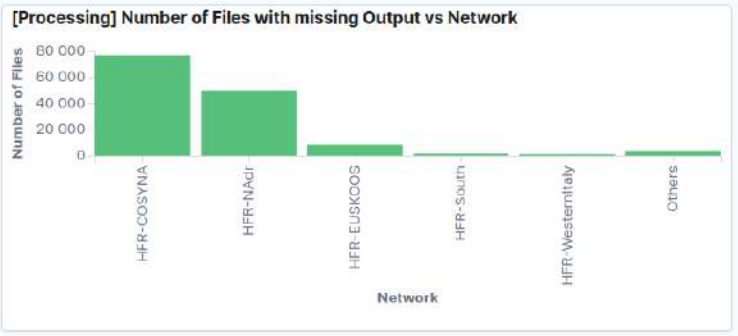


Next, plots analysing the processes based on whether they are a "Radial" or "Total" analysis and on whether we have input registry, output registry or both are shown.

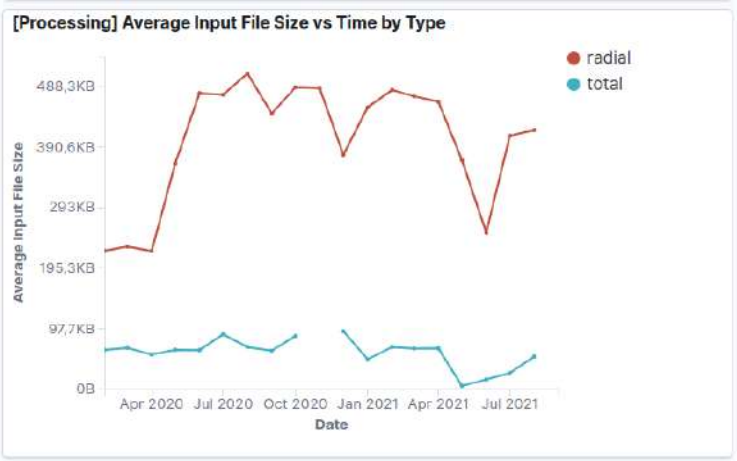


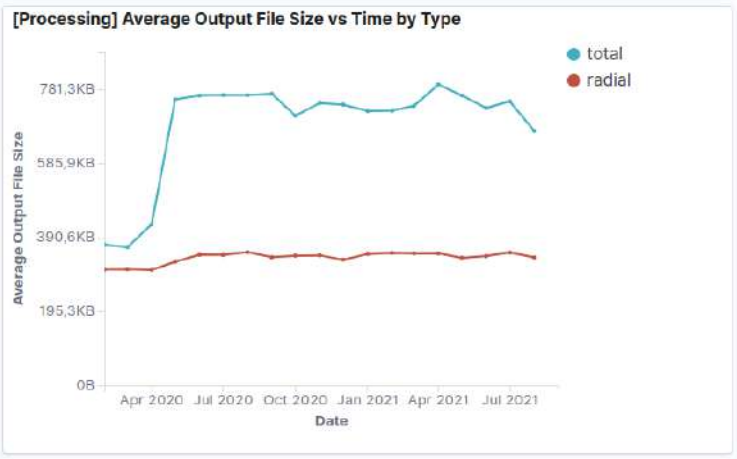


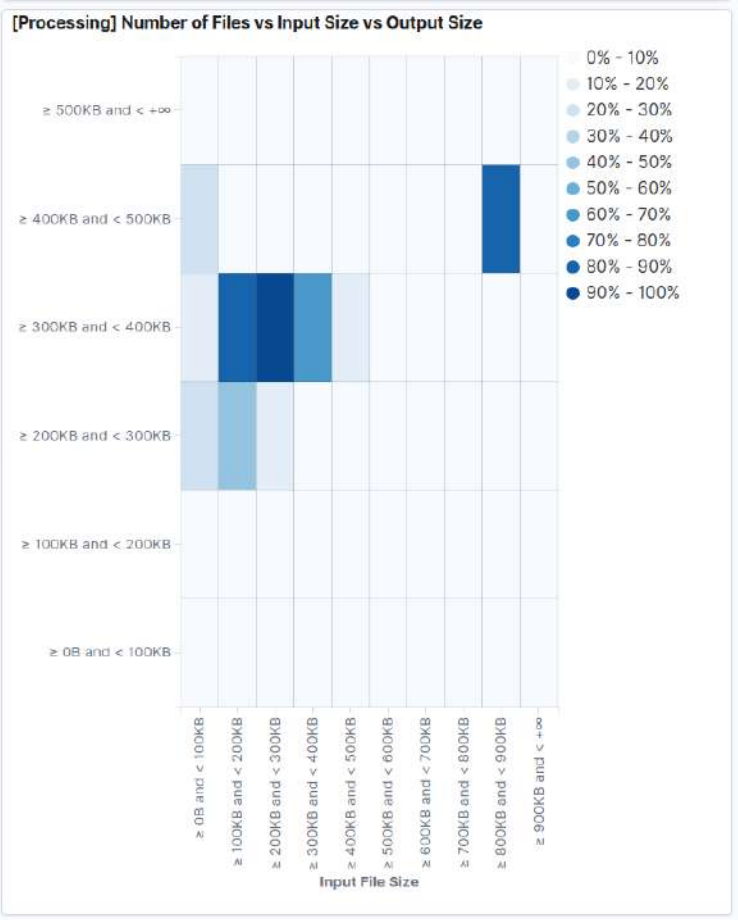




Next, plots analysing the processes based on whether they are a "Radial" or "Total" analysis and on whether the process has input registry, output registry or both are shown.





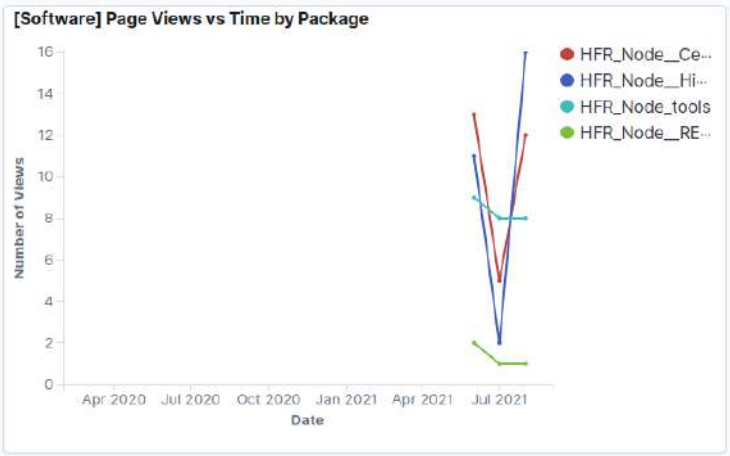


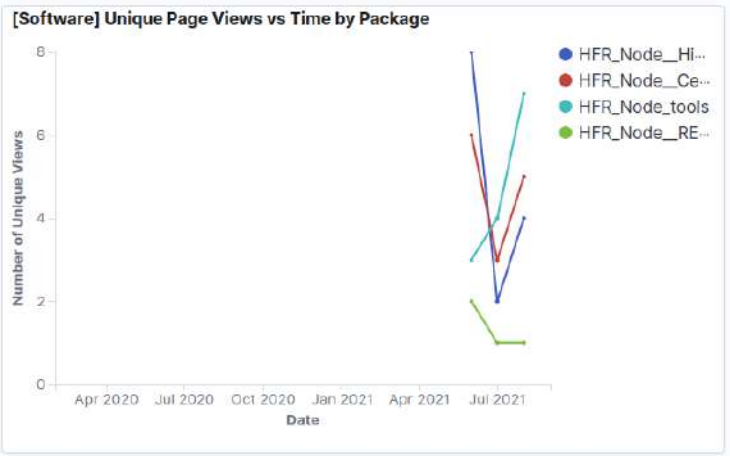
It can be noted that COSYNA represents an important HFR network by the number of files submitted. It might seem that this network needs very little process time. However, this is because many of the processes are missing the output registry. Additionally, an increase in the number of processes that miss the output registry is also worth noting.

Software

The data underlying the following plot has been obtained from Github API. Data is only available from June 18th 2021 because Github only keeps access metrics for 15 days. Therefore, data regarding access 15 days before the development of VAMS was not kept.

The following plots show the number of Views and Unique Views that each software package receives on GitHub





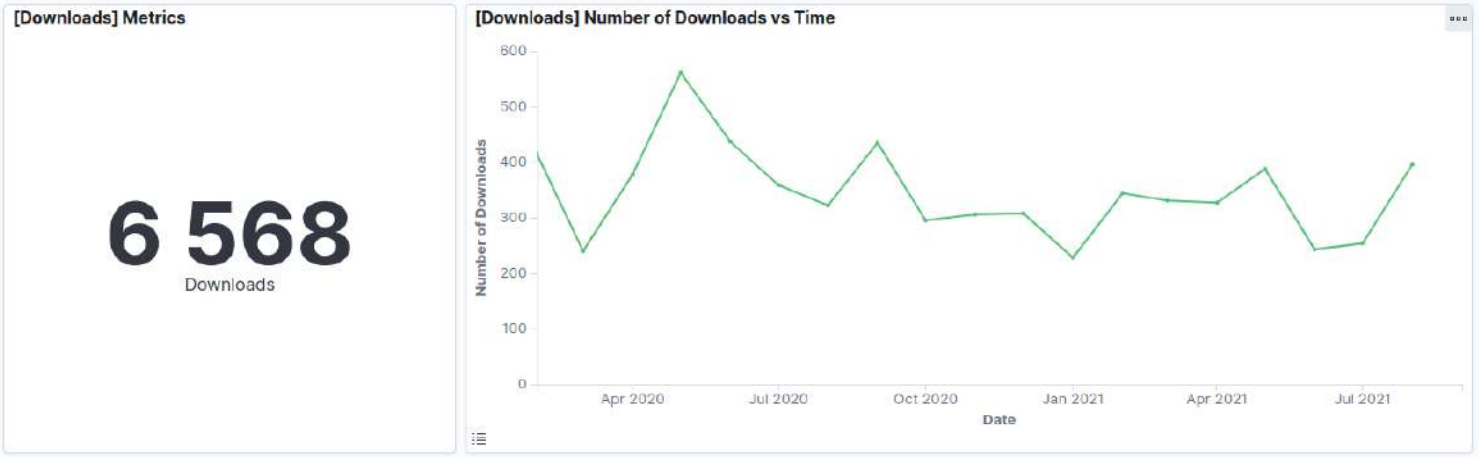
Annexe 1.6 CytoFluoTool (ID 8.1)

The access to the CytoFluoTool software package is monitored in two different ways: the CRAN downloads API and the log files of the documentation website. Next, metrics based on these two ways of monitoring will be presented.

Downloads from CRAN

The data underlying the following plots have been obtained from the CRAN downloads API.

The next plot shows the total number of downloads and their temporary distribution.



Visits to the Documentation Page

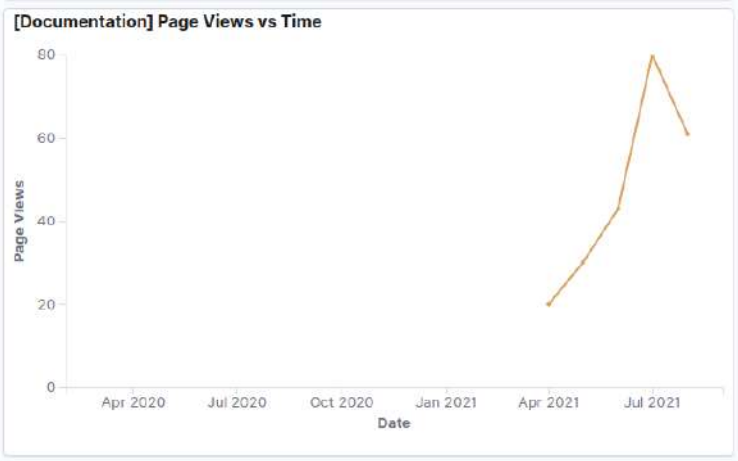
Software usage cannot only be measured by downloads, as a piece of software can be downloaded once and used many times. Therefore, It has been decided to measure access to the documentation pages too, in order to be able to estimate the trend in the usage of the software.

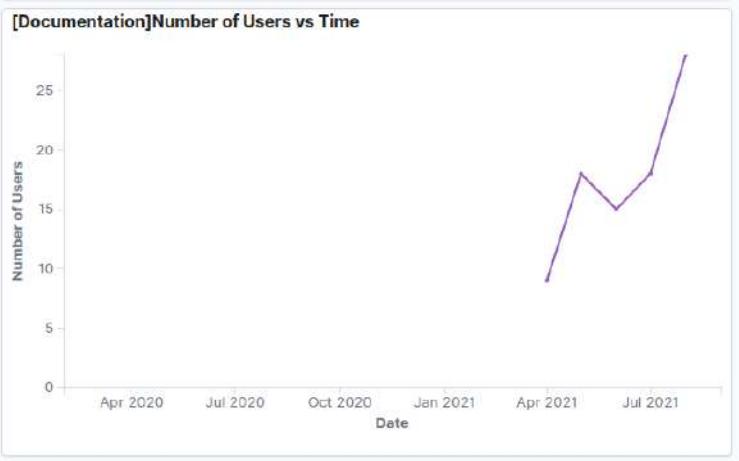
The data underlying the following plots have been obtained from the log files of the server of mawenzi's website. Unfortunately, logs of the first months of JERICO-S3 had been lost. Only data from April 13, 2021 is available.

First, the general metrics and geographic analysis of the visits will be presented. Then, the temporal evolution of the traffic will be seen. Finally, the content visited and the origin of traffic will be analysed.

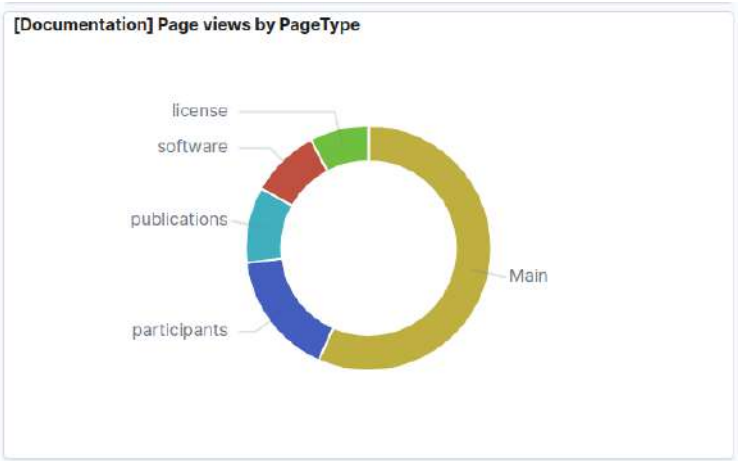
In the next image, the overall number of visits, users and countries who accessed CytoFluoTool's documentation and the geographic distribution of visits are presented.

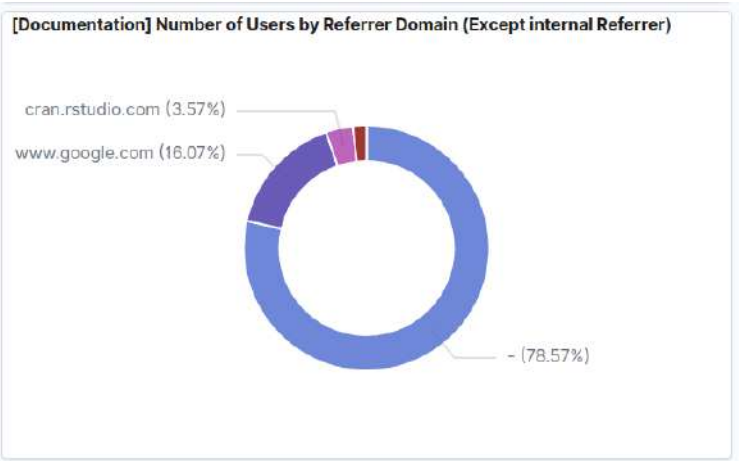
Next, two plots that show the temporal trend of pageviews and users respectively are shown.





Finally, the traffic origin is analysed: by both referrer ("-" means referrer is not informed, like a visit from a bookmarked link or direct access via software) and country.





Annexe 1.7 JERICO-ECOTAXA (ID 8.2)

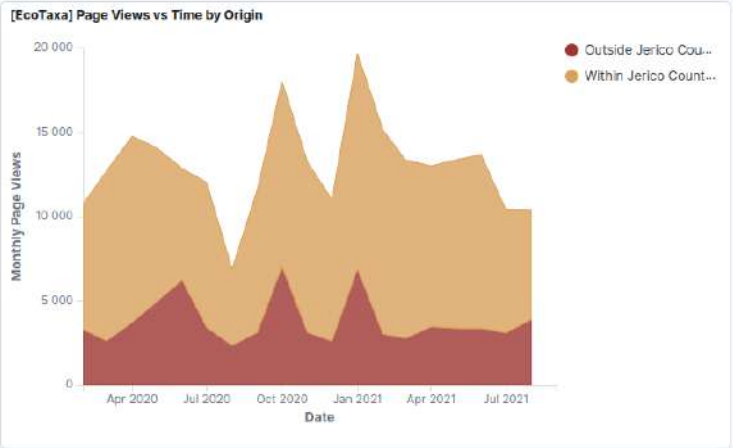
The data underlying the following plots are kindly generated by EcoTaxa's team and made available in a CSV file. There are two access modalities to EcoTaxa. One is uploading images for analysis (related images are container by the same "object", which is what we measure). The other one is browsing the website. Both of them are measured. The metrics regarding the number of objects are only available from May 13th, 2021.

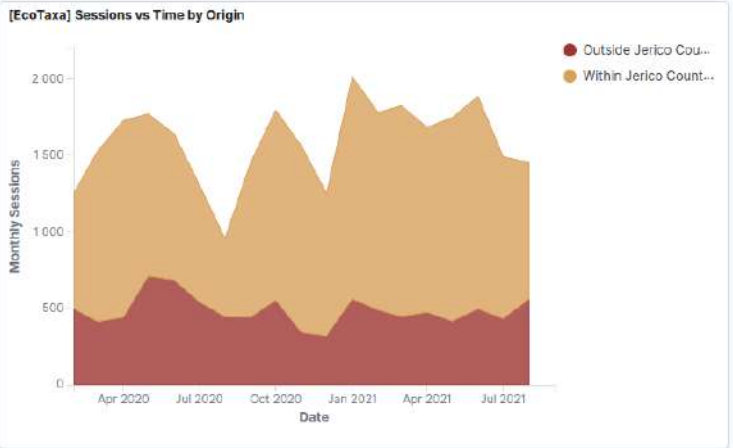
First, some overall metrics are shown. Next, analysis of objects over time, the geographic distribution of traffic and the bounce rate analysis are presented.

In the next image, a numeric representation of the access metrics is presented: insertion of objects on the left and web browsing on the right.

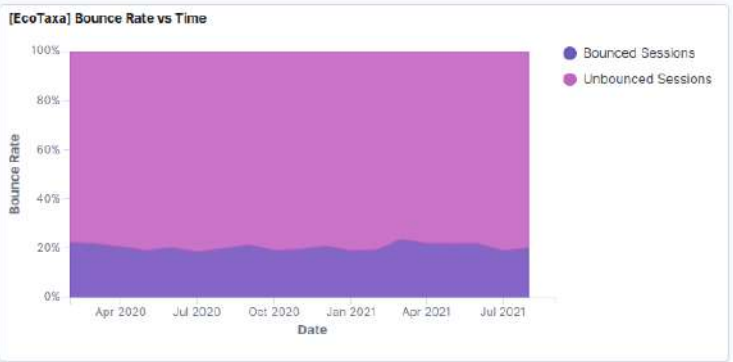
In the following image, a temporal analysis of the number of objects has been added.

Next, the sessions and pageviews from different geographic origins is analysed (countries considered to be part of JERICO for the purpose of these plots are: Belgium, Croatia, Denmark, Estonia, Finland, France, Germany, Greece, Ireland, Italy, Netherlands, Norway, Portugal, Spain, Sweden, United Kingdom).





Finally, acknowledgements to the main funders of EcoTaxa and bounce rate analysis are shown.



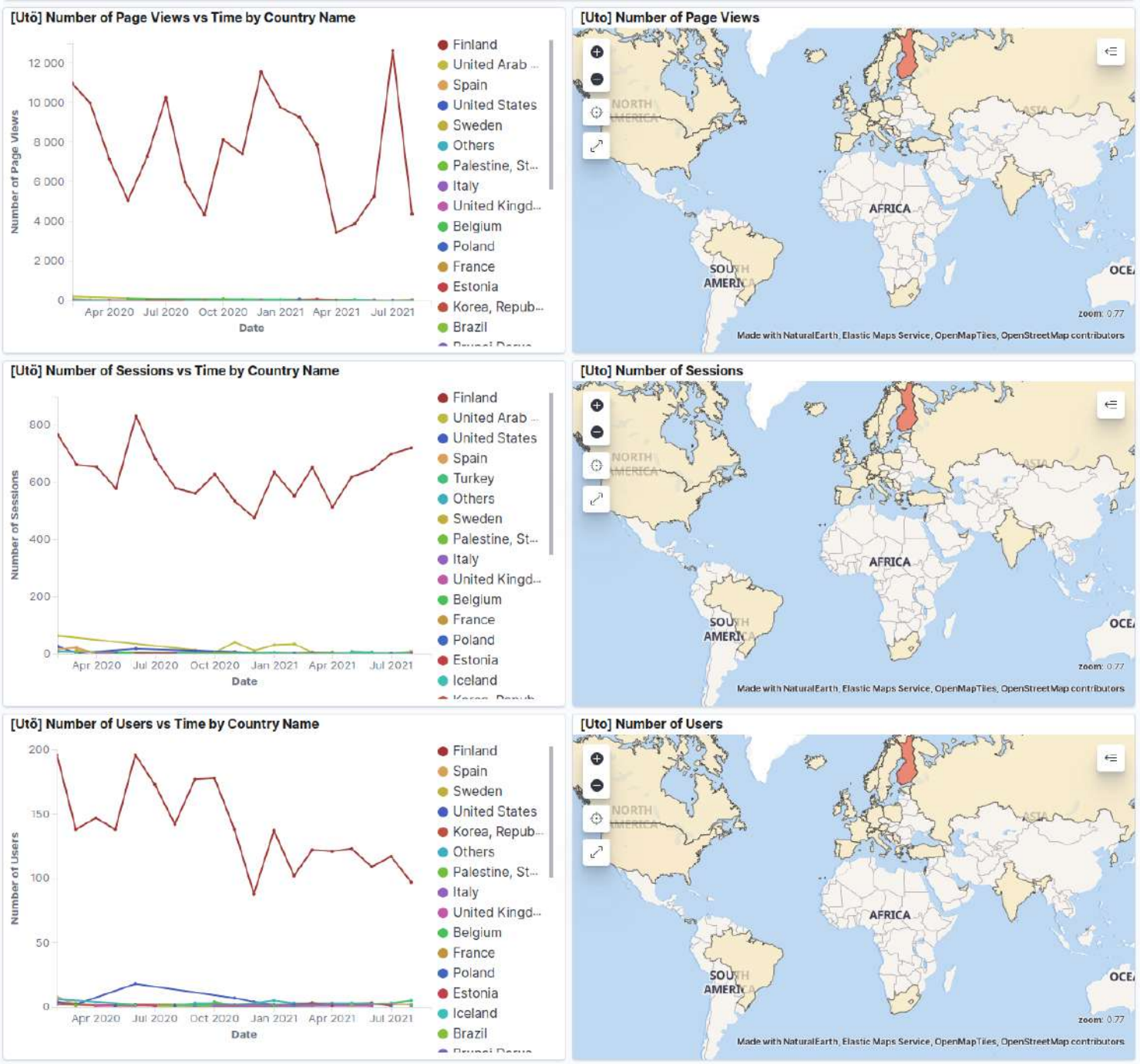
Annexe 1.8 Utö Atmospheric and Marine Research Station (ID 15.1)

The data underlying the following plots have been obtained from the Google Analytics API.

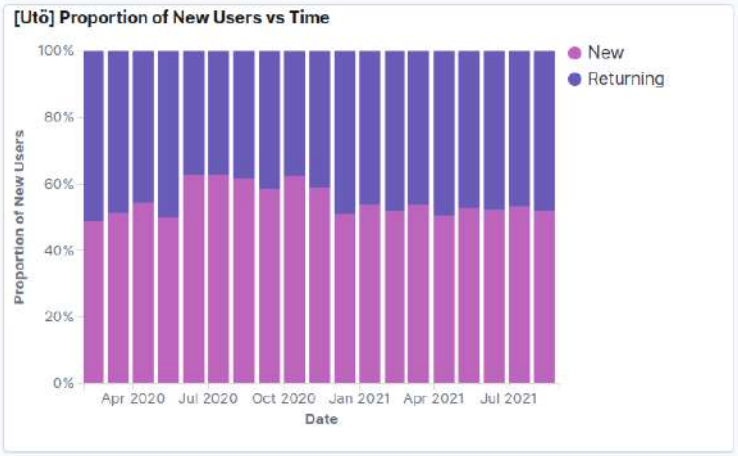
First, the overall access metrics will be presented. Then, the following analysis are included: geographic analysis, analysis by user type (new vs returning), bounce rate analysis, user device analysis and traffic source analysis.

In the next image, the general access metrics in a numeric form are presented.

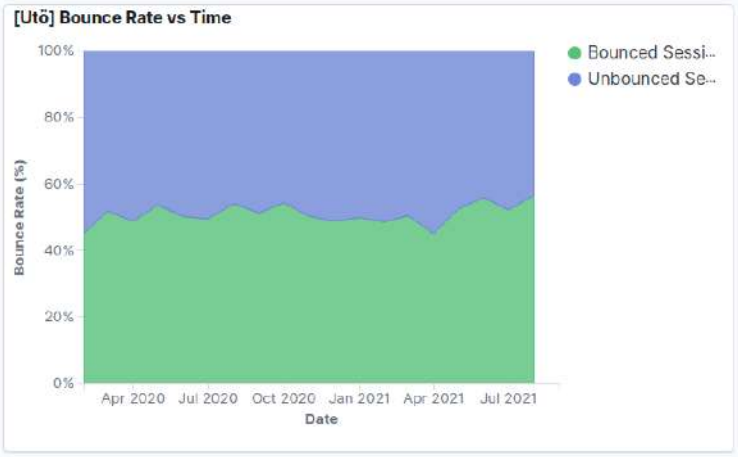
The following image contains the geographic analysis of the traffic.

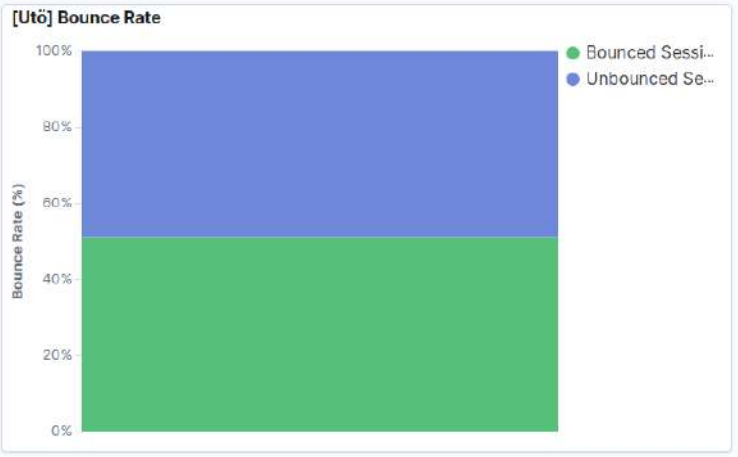


Next, plots analysing the proportion of new users are shown.

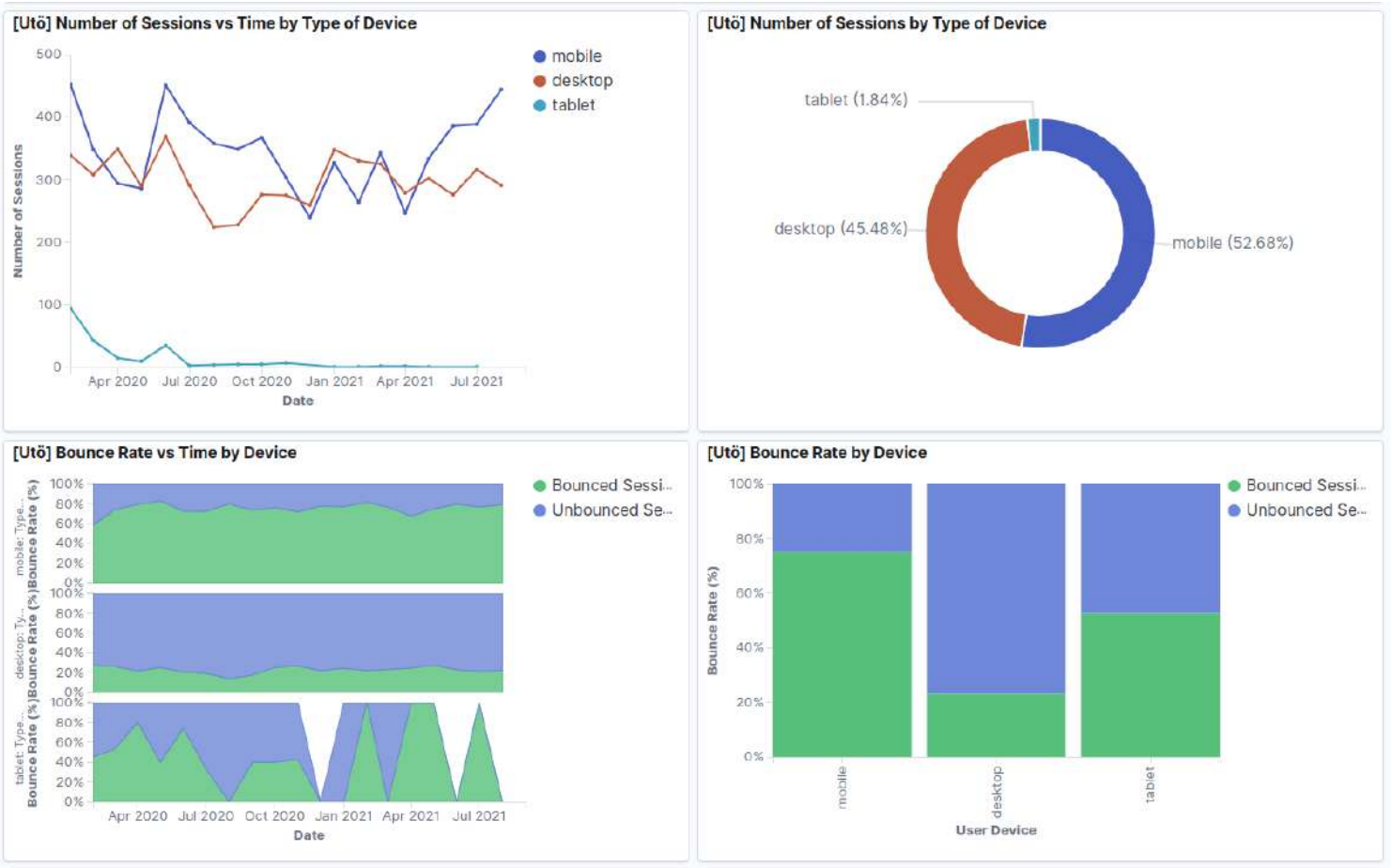


Next, the bounce rate is analysed.

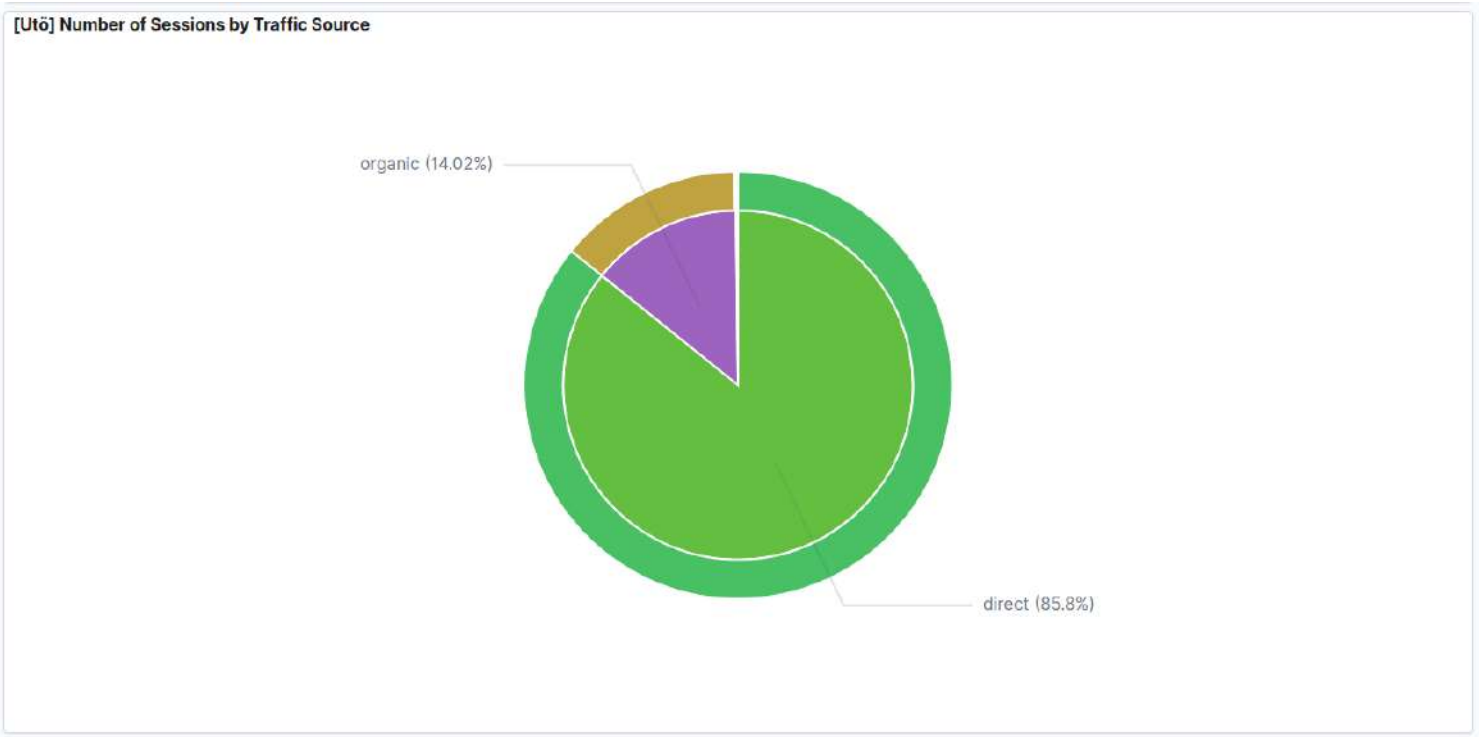




The image below shows the proportion of sessions that come from each type of device and the influence of the type of device on the bounce rate.



Finally, an analysis of traffic based on the source is presented.



Annexe 1.9 POSEIDON Multi Platform Observatory Data Center (ID 16.1)

This VA infrastructure consists of two parts: an API that is monitored by analysing the server's log files and a website monitored via Google Analytics. Next, the access metrics of each of these parts are presented.

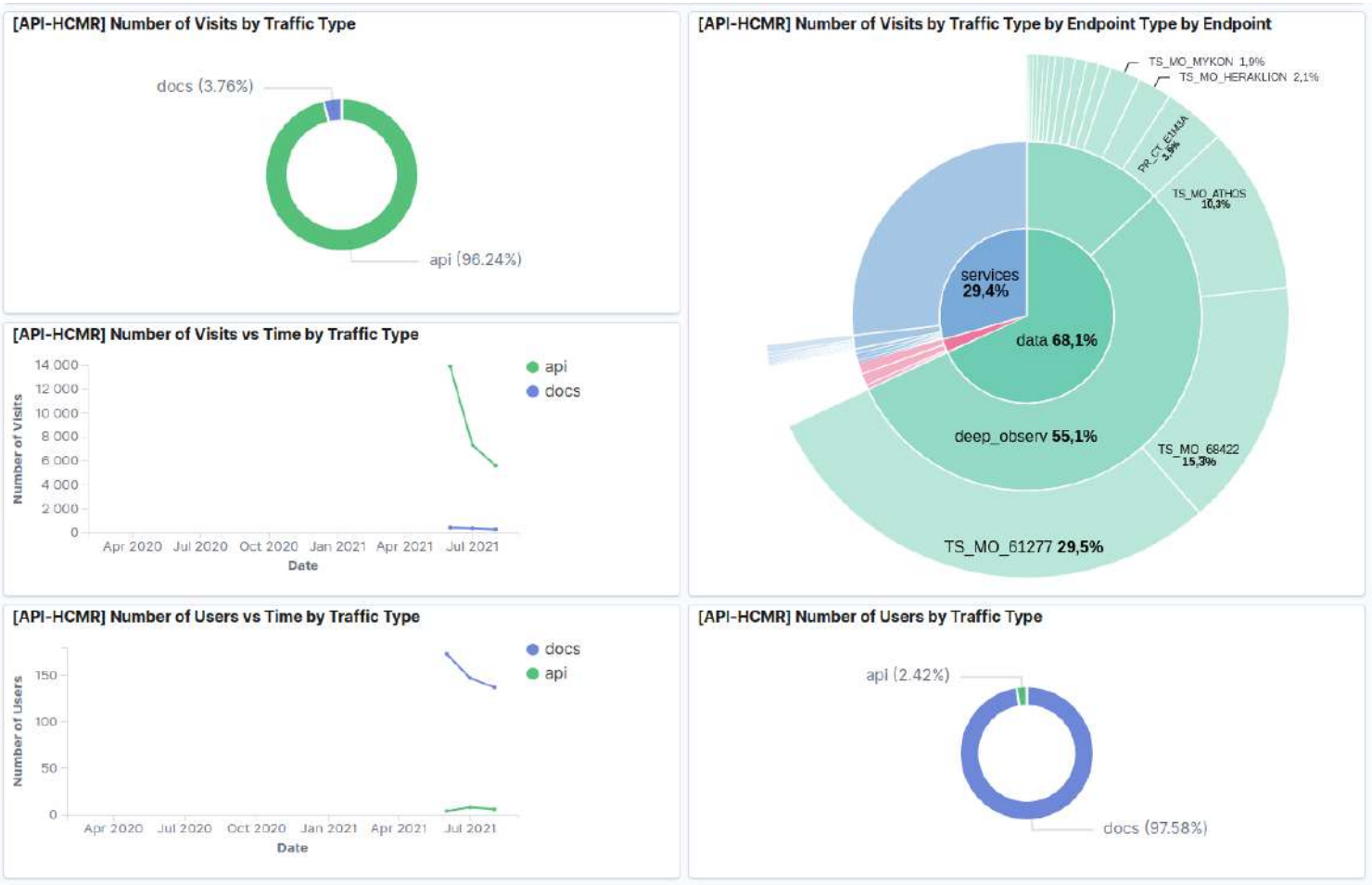
API

The data underlying the following plots have been obtained from the log files of the API server. This API was launched on June 6th 2021. So, there is no data previous to that date.

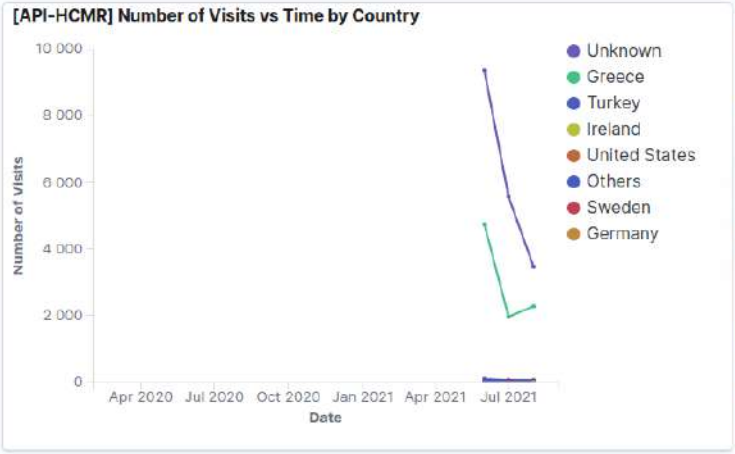
First, the overall access metrics will be presented. Then, traffic will be analysed by content, geographic origin, used client, referrer domain and operating system.

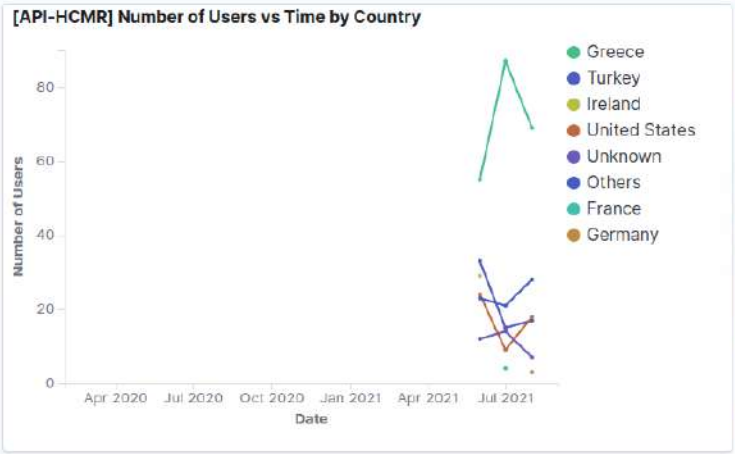
In the next image, the general access metrics in a numeric form are presented.

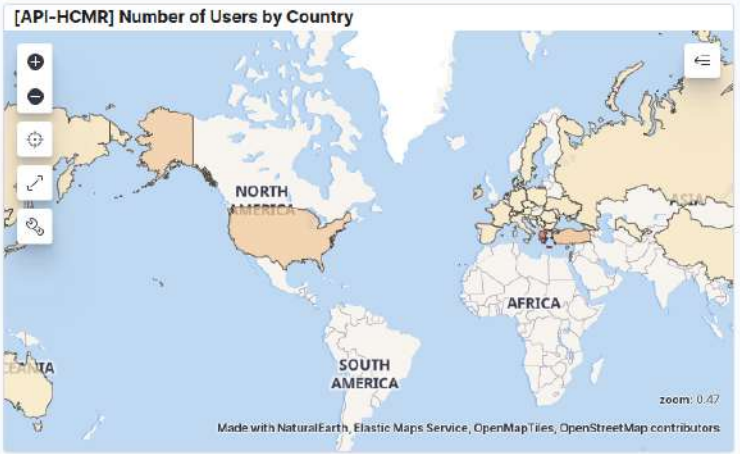
In the following image, the content is analysed. The proportion of users and pageviews that go to documentation pages or API endpoints can be seen. It is also possible to see how the traffic to API endpoints is distributed among the three kinds of endpoints (data, services or metadata), what specific endpoints are being visited and for what platform (when the endpoint provides information from one specific platform).



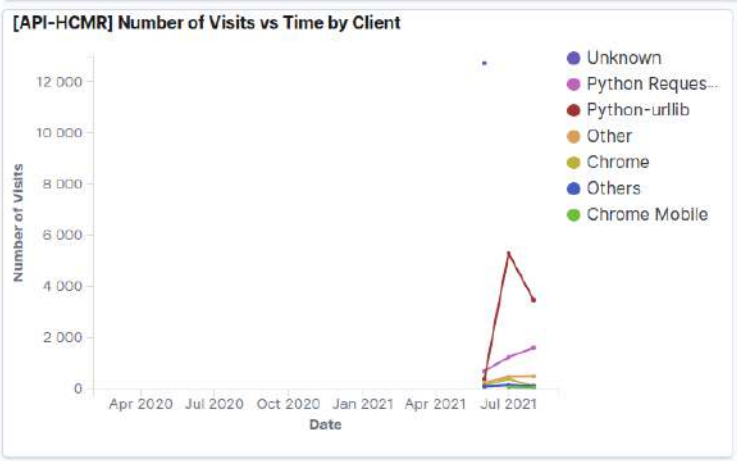
Next, plots and maps with a geographic analysis of the traffic seen by the server are shown.

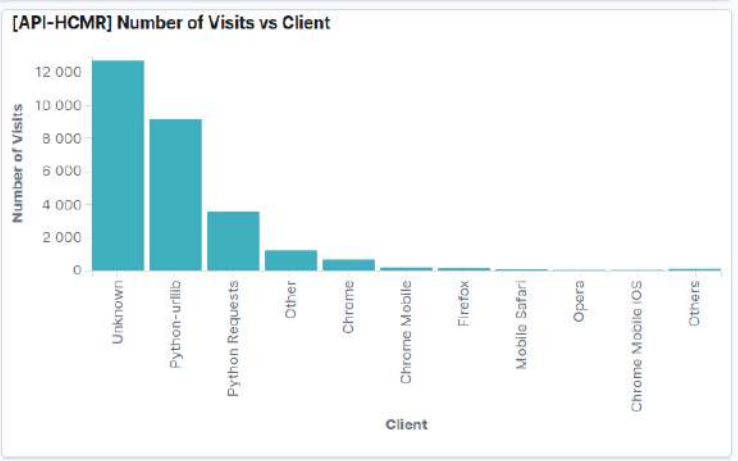


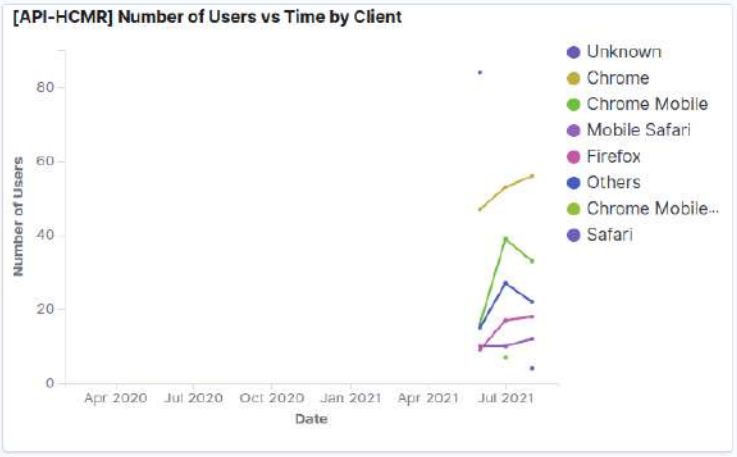


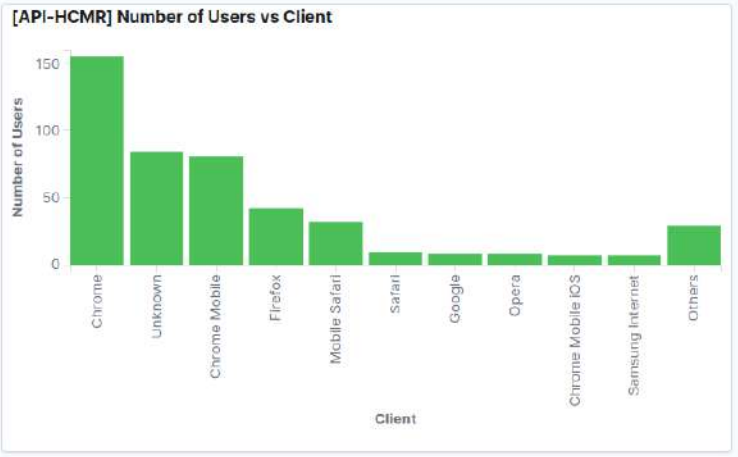


The fourth image shows an analysis of the traffic based on the software used by the user to perform the request (a browser or an HTTP library).

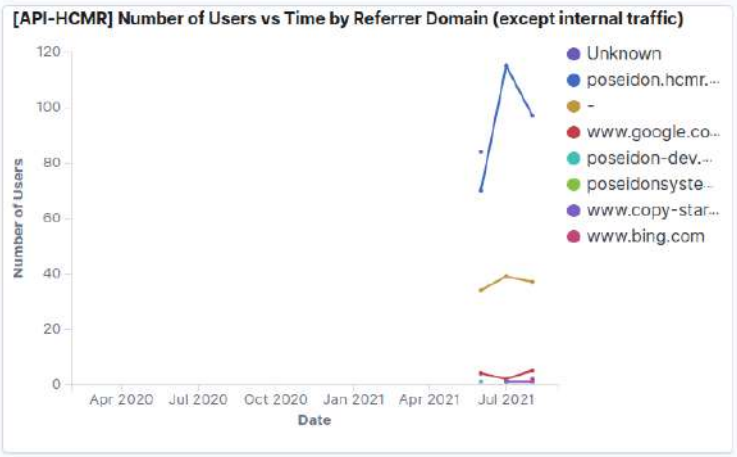


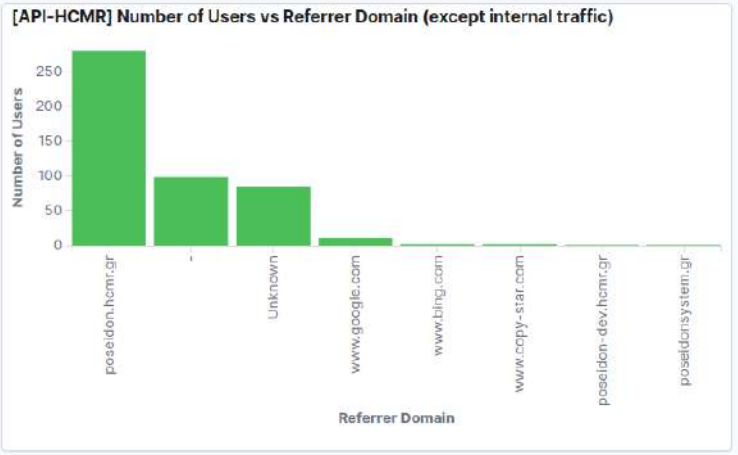






The image below contains an analysis based on the referrer domain ("-" means referrer is not informed, like a visit from a bookmarked link or direct access via software).





Finally, an analysis of traffic based on the operating system of the user is presented (that information is extracted from the browser version, for HTTP libraries, the information is not available and it is contained in the "Other" category).

A difference in the number of visits between June 2021 and the rest of the months can be seen. This difference is concentrated in the API endpoints as opposed to the documentation pages and cannot be seen in the number of users. This means that one or more users did extensive use of the API endpoints in the first month of operation, which suggests an intense testing phase. Due to the short time series, it is not possible to see longer-term trends.

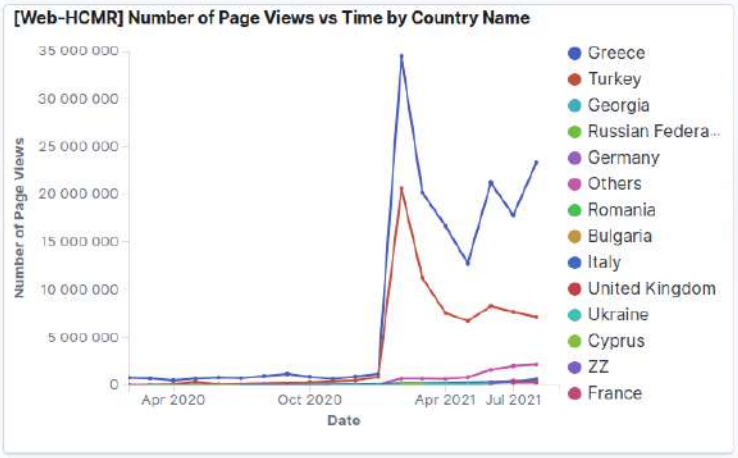
Website

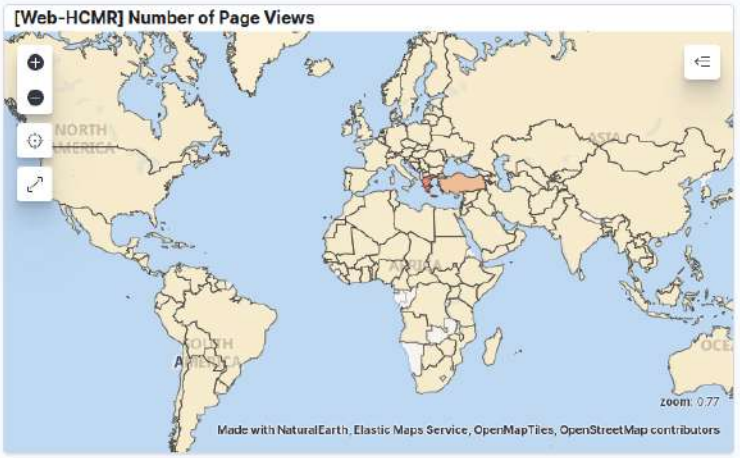
The data underlying the following plots have been obtained from the Google Analytics API.

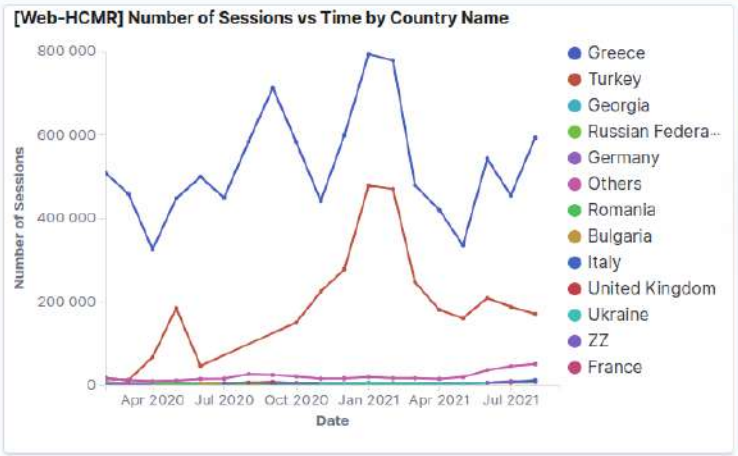
First, the overall access metrics will be presented. Then, the following analysis are included: geographic analysis, analysis by user type (new vs returning), bounce rate analysis, user device analysis and traffic source analysis.

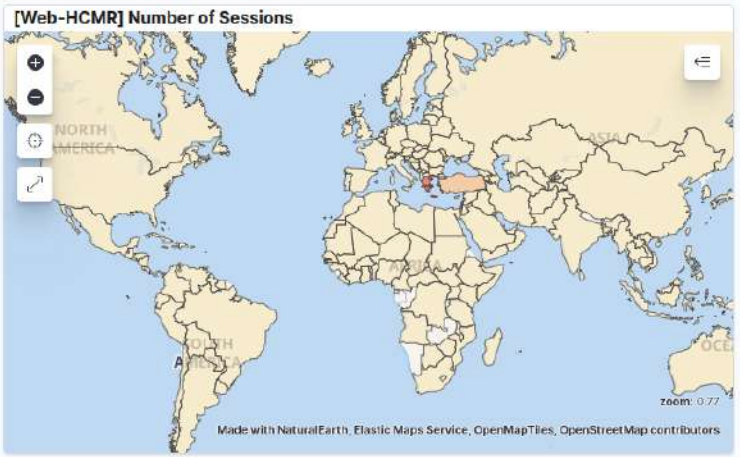
In the next image, the general access metrics in a numeric form are presented.

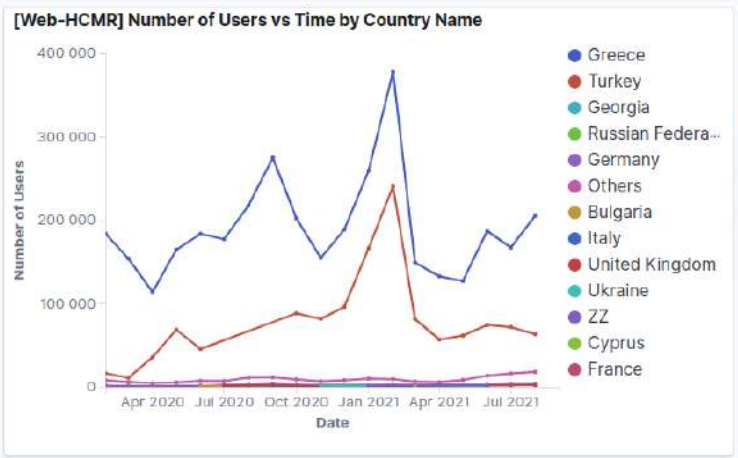
The following image contains the geographic analysis of the traffic.

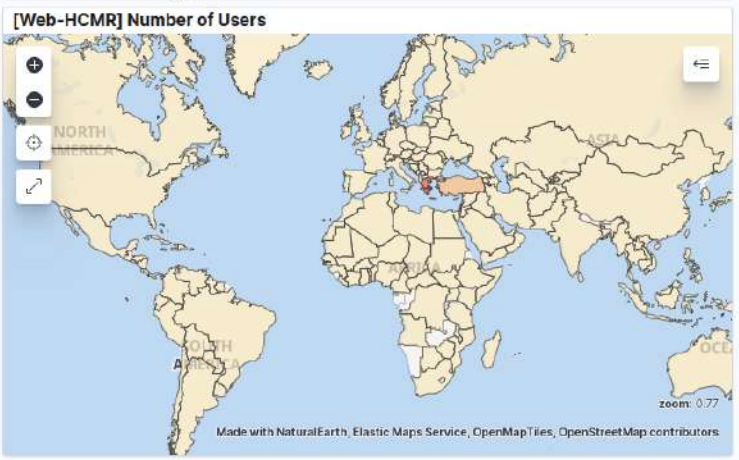




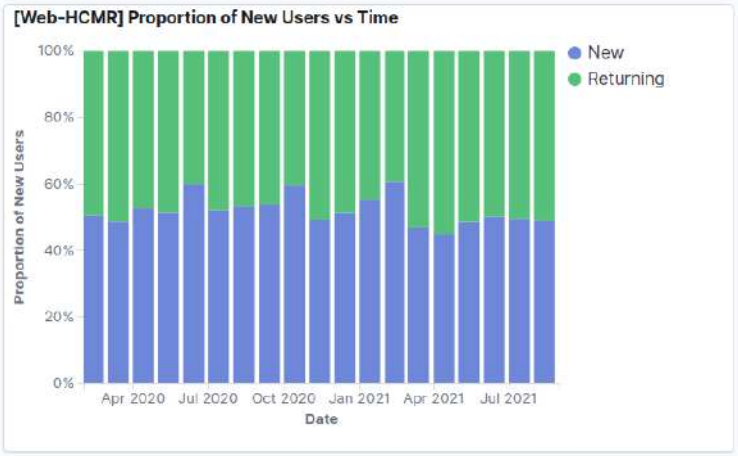


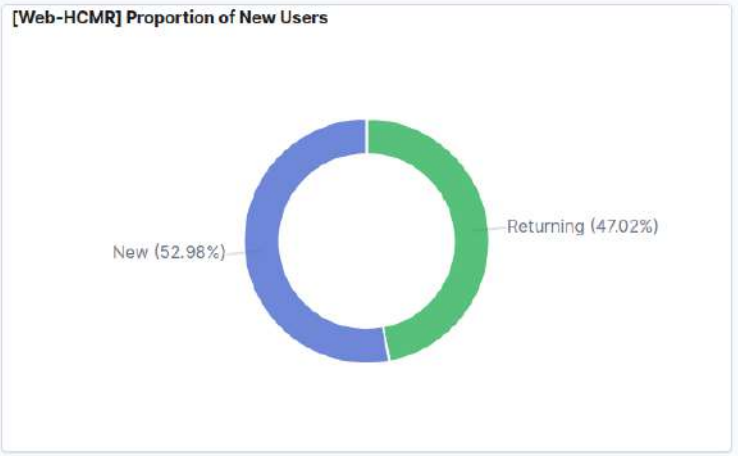




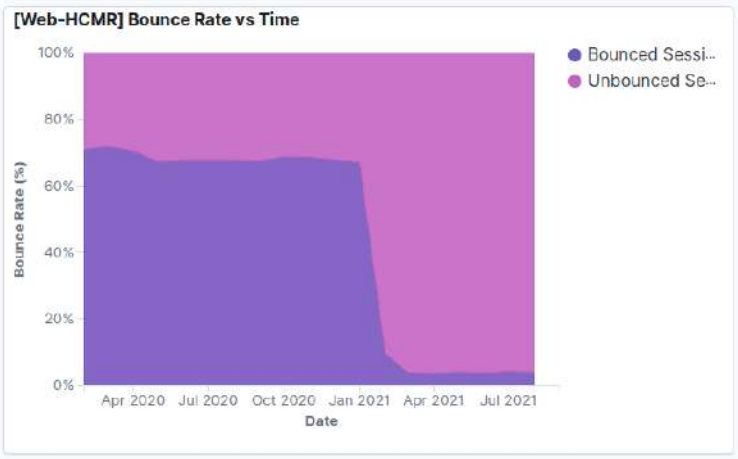


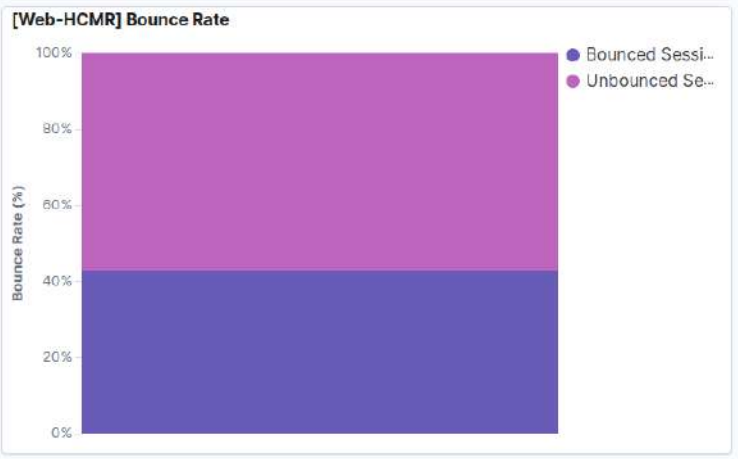
Next, plots analysing the proportion of new users are shown.



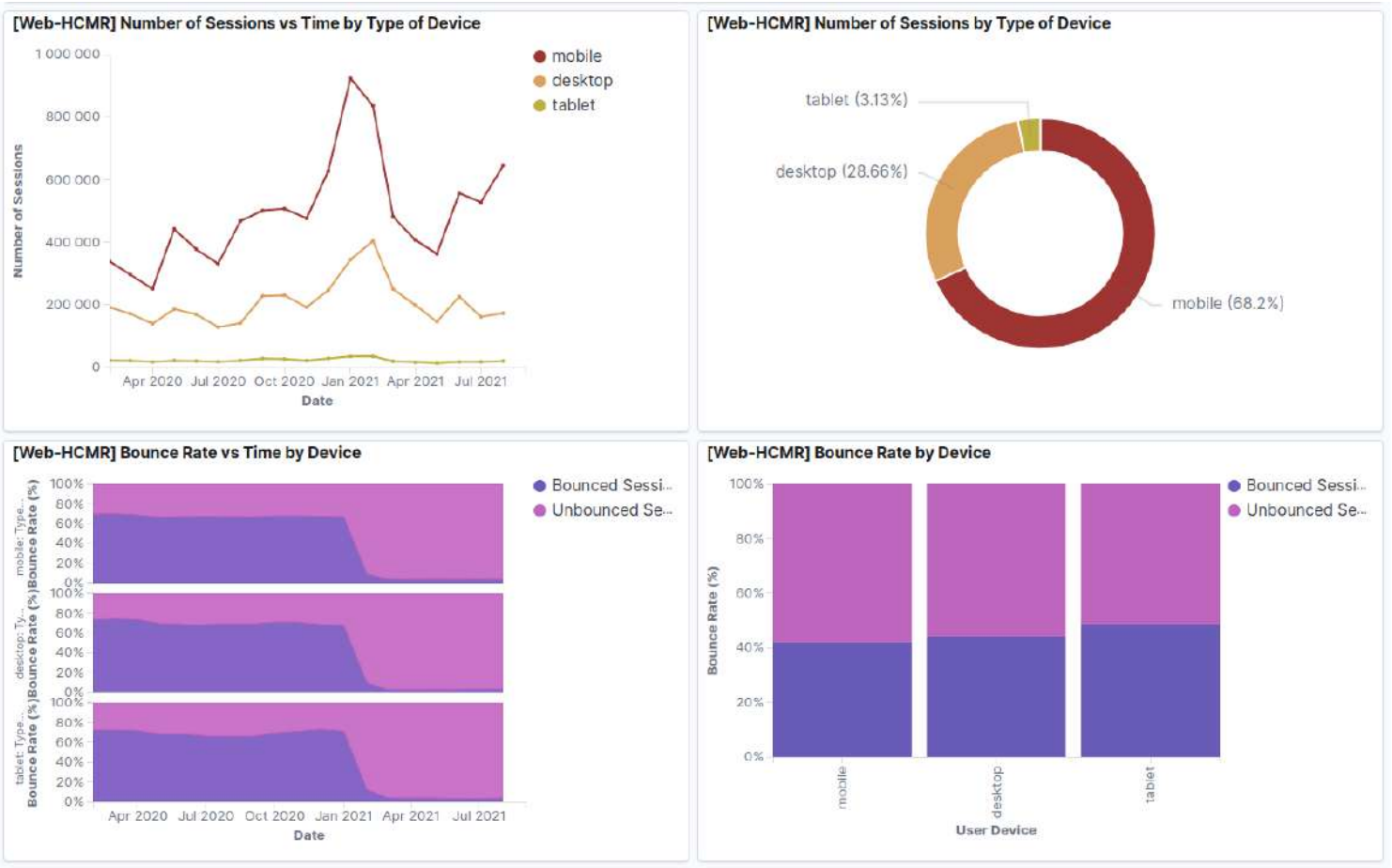


Next, the bounce rate is analysed.

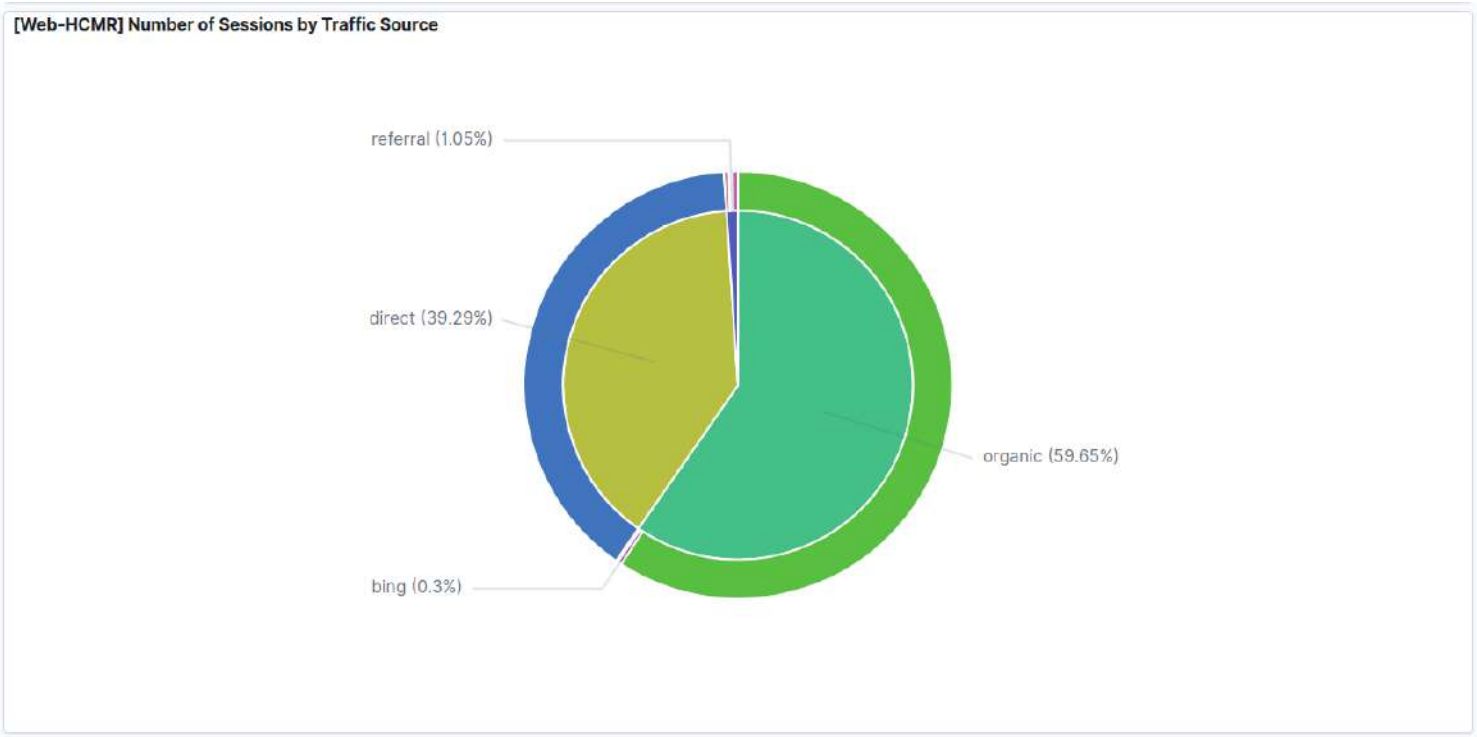




The image below shows the proportion of sessions that come from each type of device and the influence of the type of device on the bounce rate.



Finally, an analysis of traffic based on the source is presented.



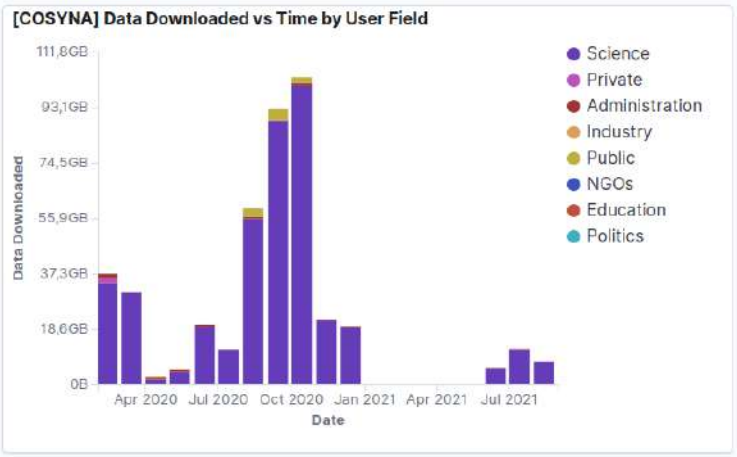
There is a clear change in user behaviour after January 2021 which correlates with the release of the new Poseidon website.

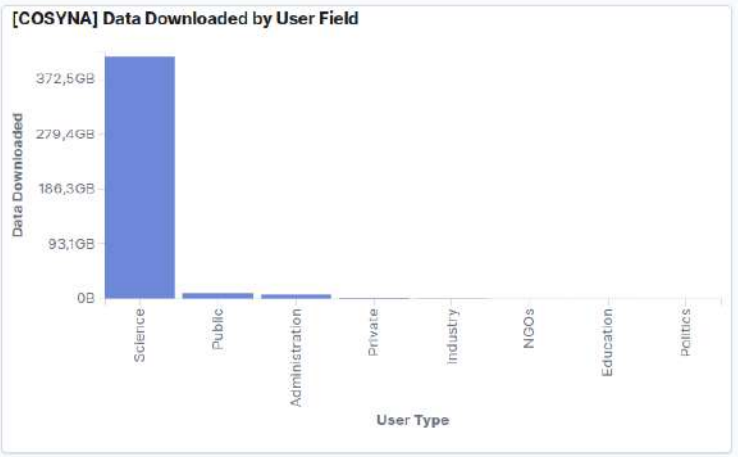
Annexe 1.10 The Coastal Observing System for Northern and Arctic Seas (COSYNA) (ID 17.1)

The data underlying the following plots are obtained via a Google Form that COSYNA staff kindly fill monthly based on their internal monitoring system. Due to a technical issue in that system, there is no data available between January 2021 and May 2021 (June 2021 has partial data only).

First, the traffic by the type of user will be analysed. Then, the analyses will be done by origin country.

In the next image, it can be seen how the traffic is distributed among the types of users.





In the following image, the proportion of data that is used by users from COSYNA's country (Germany) can be seen.

Overall, since February 2020, there were 1658 sessions from 297 different users. All sessions require an interactive login process. This means that only interested users are counted and accidental calls for the website or robots are not.

Annexe 1.11 HIDROGRAFICO+ (ID 19.1)

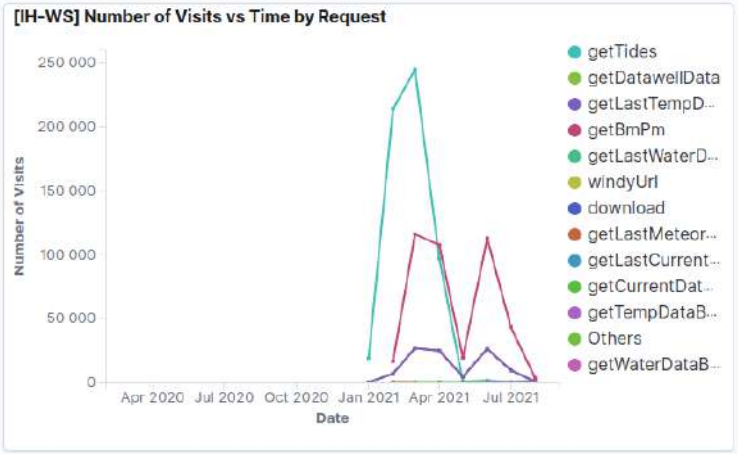
This VA infrastructure consists of two parts: a web service that provides oceanographic data, which is monitored by an internal system, and a website that helps users visualise the data, which is monitored via Google Analytics. Next, the access metrics of each of these parts are presented.

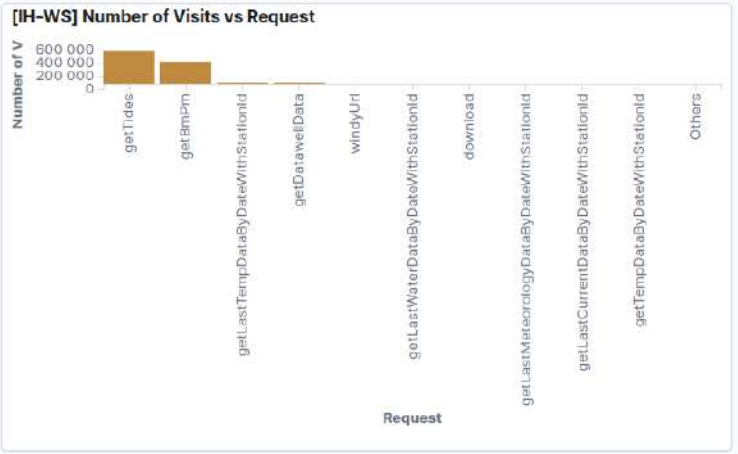
Web Service

The data underlying the following plots have been obtained by synchronizing VAMS's ElasticSearch database with IH's ElasticSearch database which is fed by an internal monitoring system. Only data since September 1st, 2021 is available.

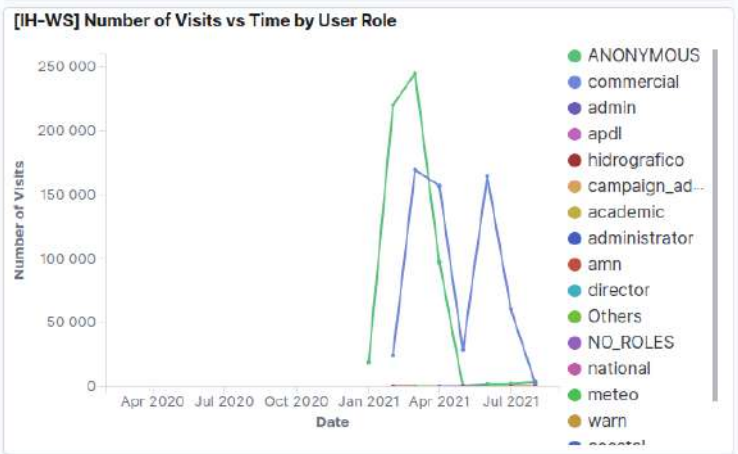
Images with pairs of plots that analyse the number of events over time and overall respectively are presented. Each image divides the events by request, user role, station id and group respectively.

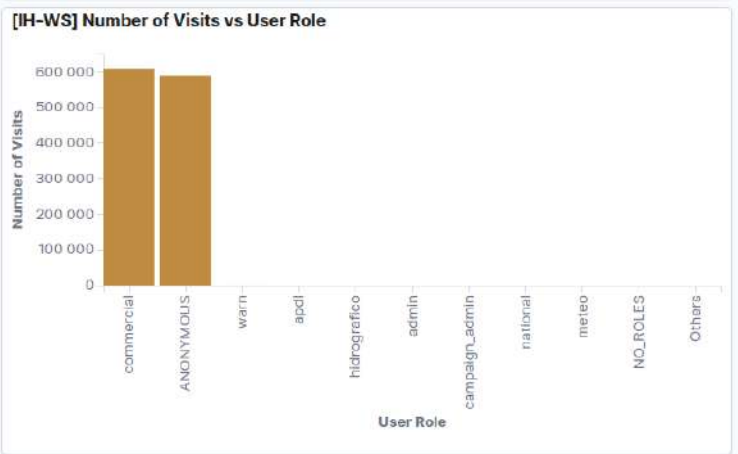
In the next image, the number of events by type of request is analysed.



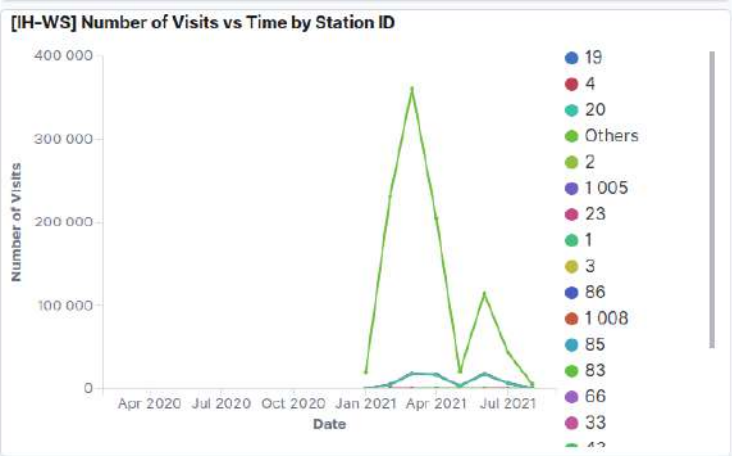


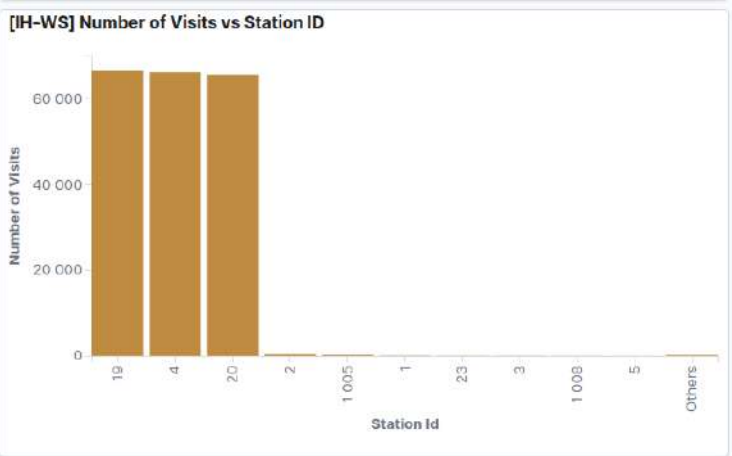
In the following image, the number of events by user role is analysed.



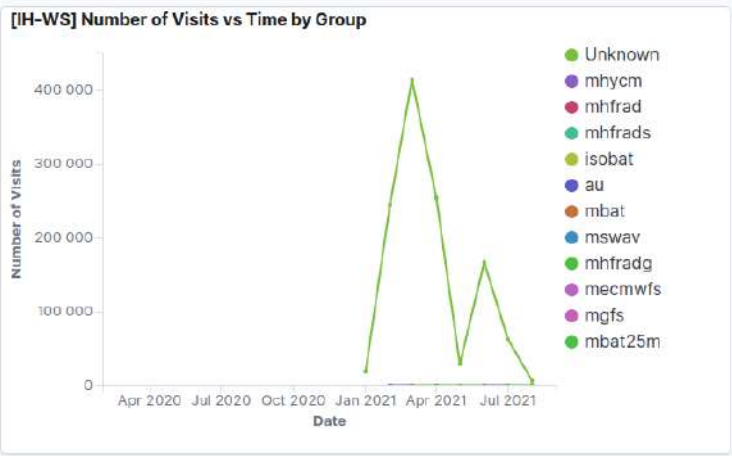


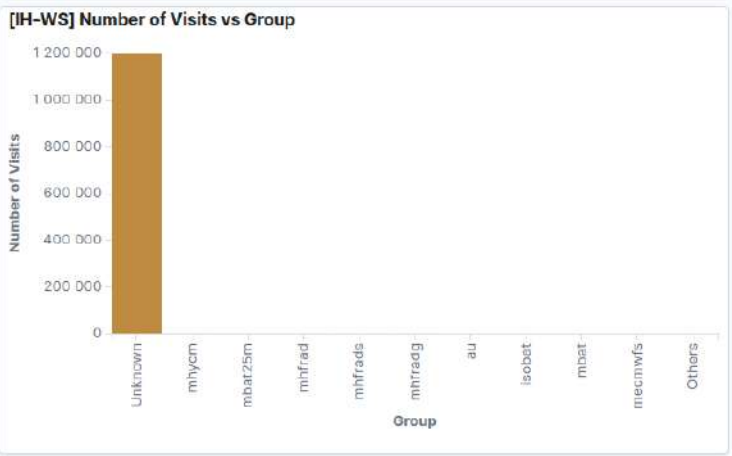
In the image below, the number of events by station Id is analysed.





Finally, the number of events by group is analysed.





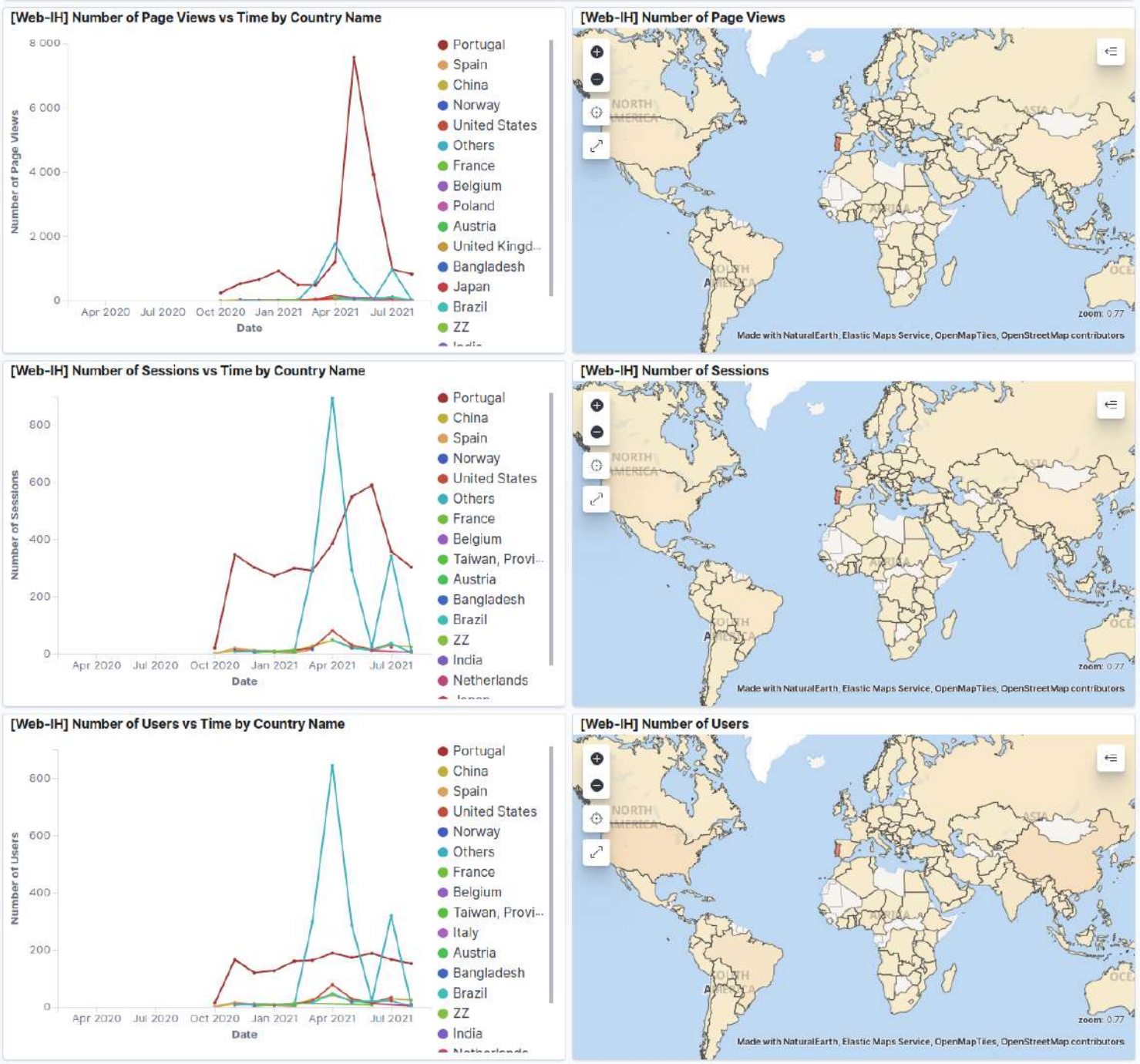
Website

The data underlying the following plots have been obtained from the Google Analytics API.

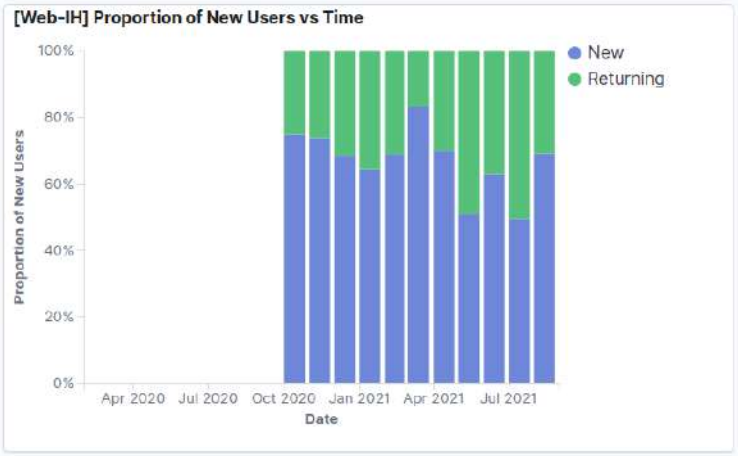
First, the overall access metrics will be presented. Then, the following analysis are included: geographic analysis, analysis by user type (new vs returning), bounce rate analysis, user device analysis and traffic source analysis.

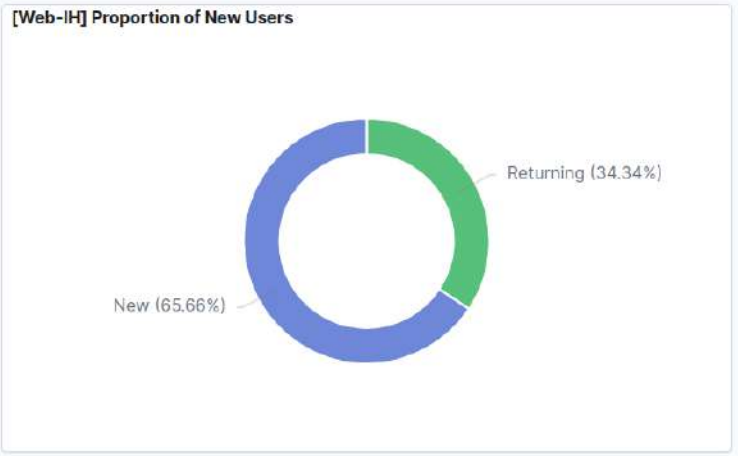
In the next image, the general access metrics in a numeric form are presented.

The following image contains the geographic analysis of the traffic.

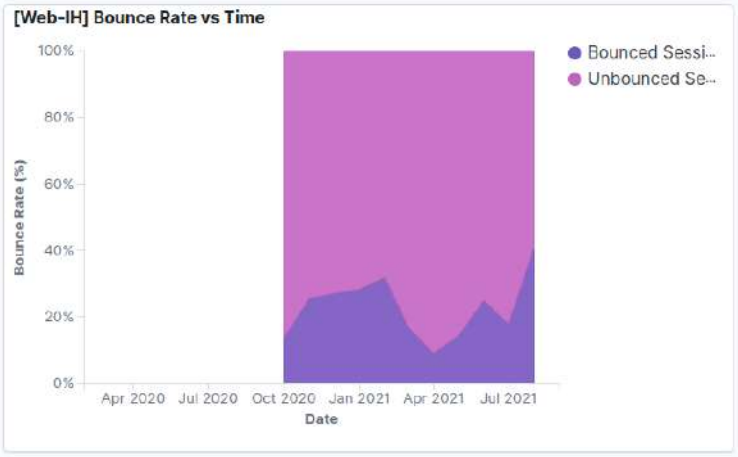


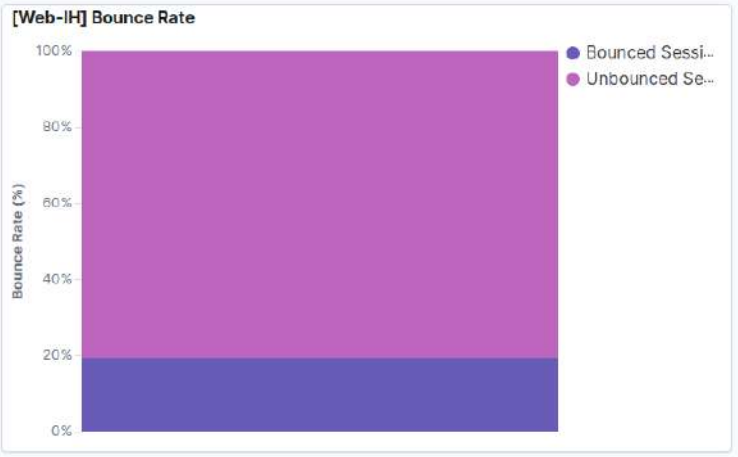
Next, plots analysing the proportion of new users are shown.



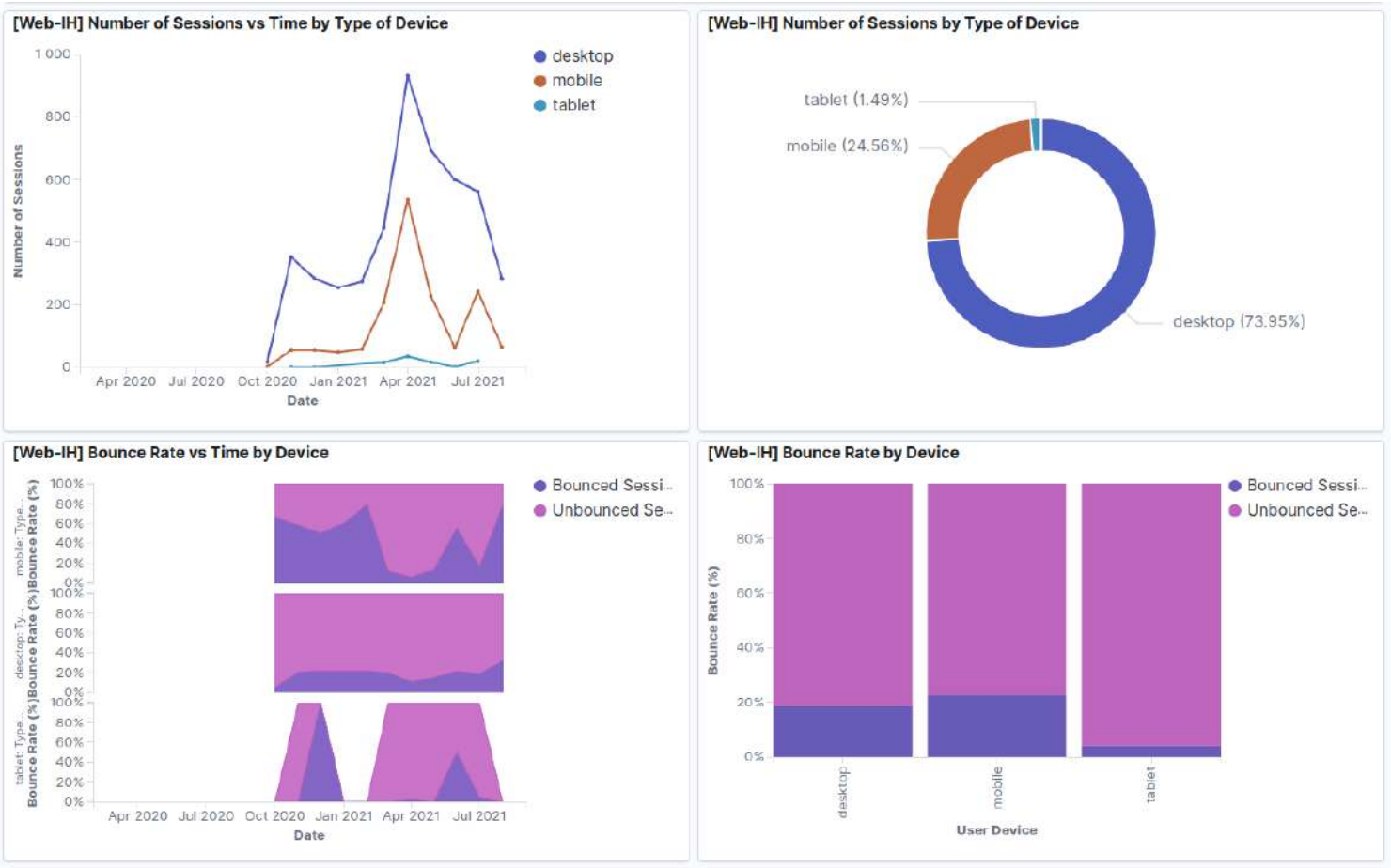


Next, the bounce rate is analysed.

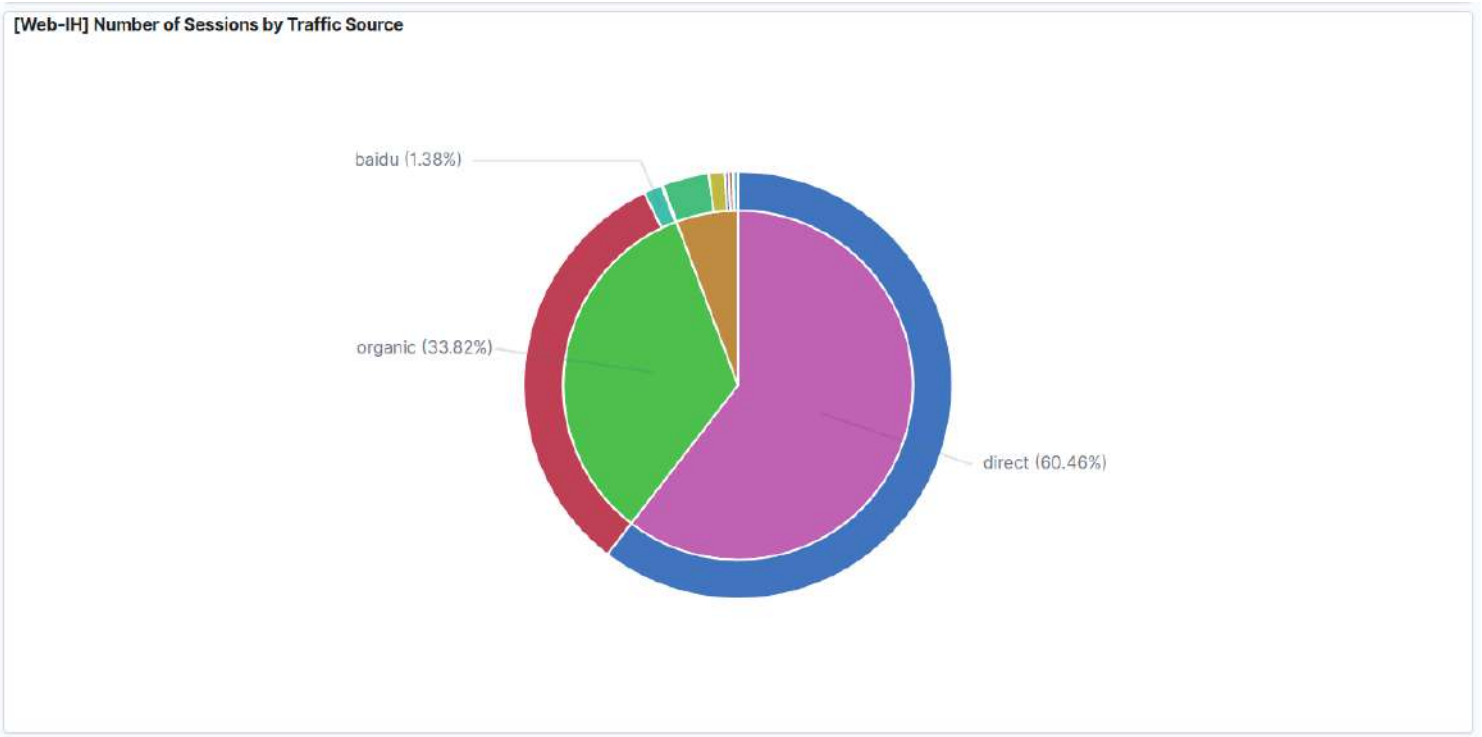




The image below shows the proportion of sessions that come from each type of device and the influence of the type of device on the bounce rate.



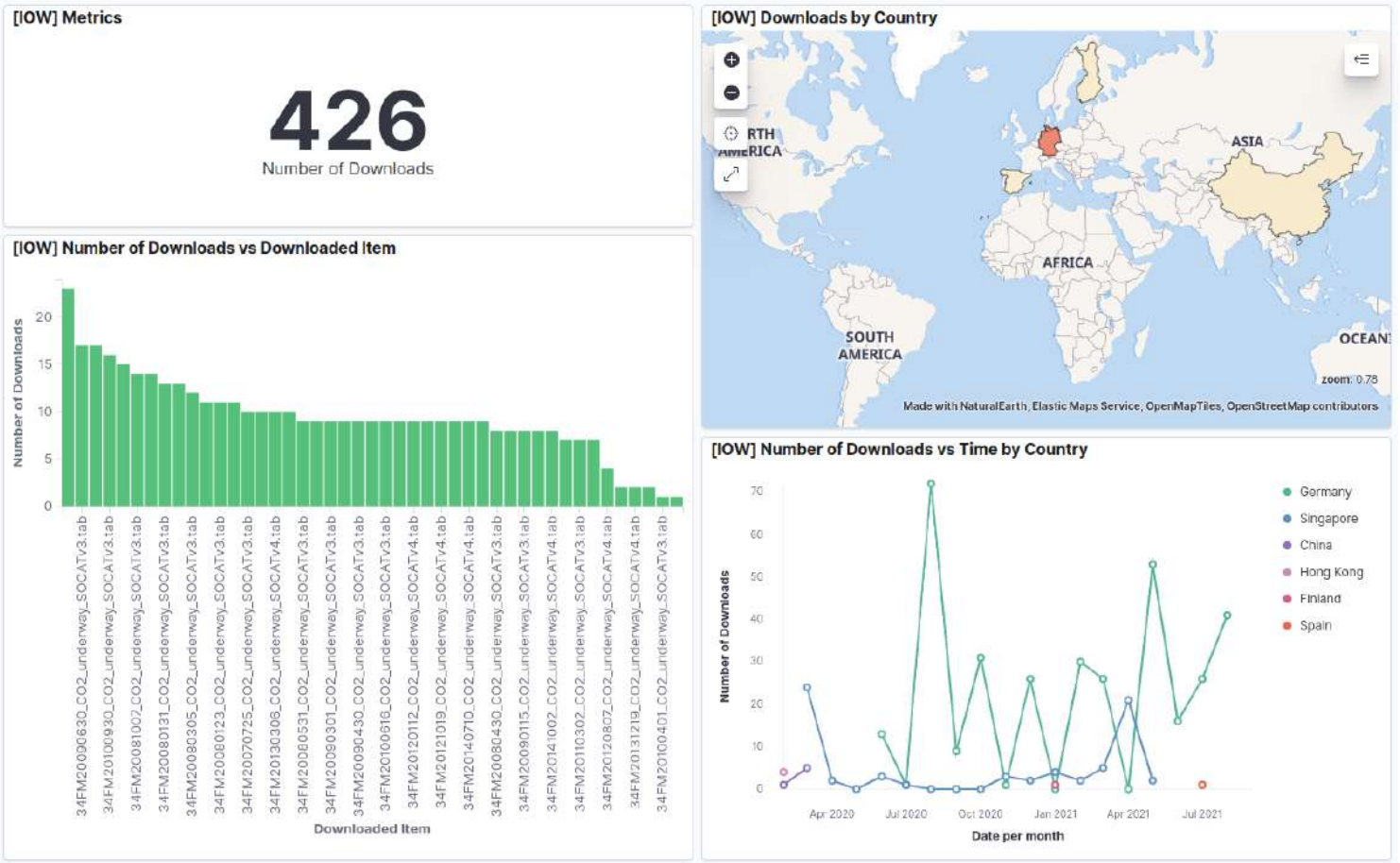
Finally, an analysis of traffic based on the source is presented.



Annexe 1.12 VOS Finnmaid GHG - BGC (ID 21.1)

The data underlying the following plots are obtained via SOCAT's Access Metrics API which SOCAT's team gracefully updated to accommodate VAMS

The following image includes the metrics of the downloads of the Finnmaid-generated files by country, file name and date.



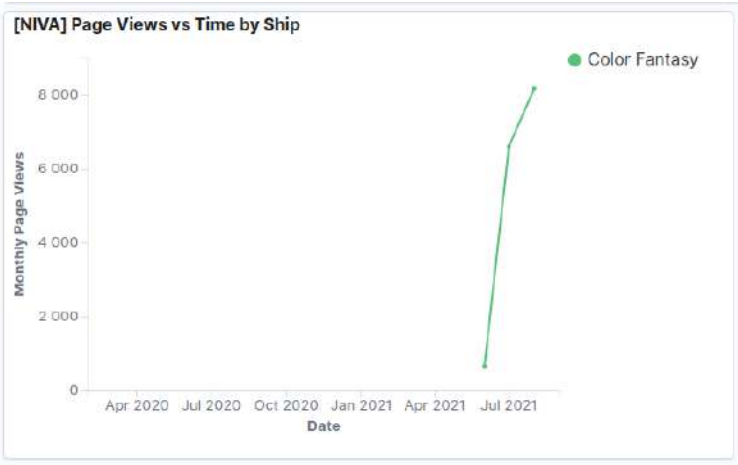
Annexe 1.13 NorFerry/NorSOOP (ID 25.1)

The data underlying the following plots are obtained via email in data files that NIVA's team sends. These files are obtained from the devices users access, which are installed on the ships. It is necessary to access a shore-based server to obtain these files due to the security policy aboard. Due to COVID, only one ship has been in operation with passengers at the time this document is published. Therefore, the data is totally incomplete and is added here only for completeness.

First, access metrics in a numeric form are presented. The rest of the images include plots that analyse the number of pageviews, the number of users and the number of days the ship has had users respectively.

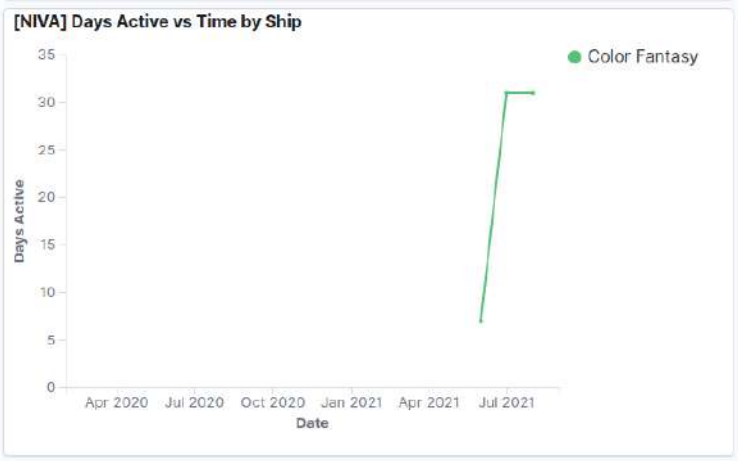
First, the access metrics are presented in a numeric form.

Next, the page views over time and overall by ship are presented.



In the following image, the number of visitors over time and overall by ship are presented.

Next, the number of active days over time and overall by ship are presented.

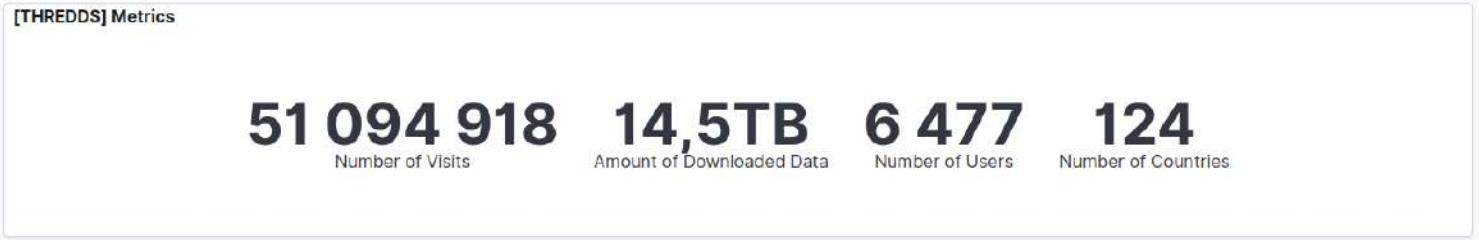


Annexe 1.14 PORTUS Observing and Forecasting System (ID 28.1)

The data underlying the following plots have been obtained from the log files of the THREDDS data server.

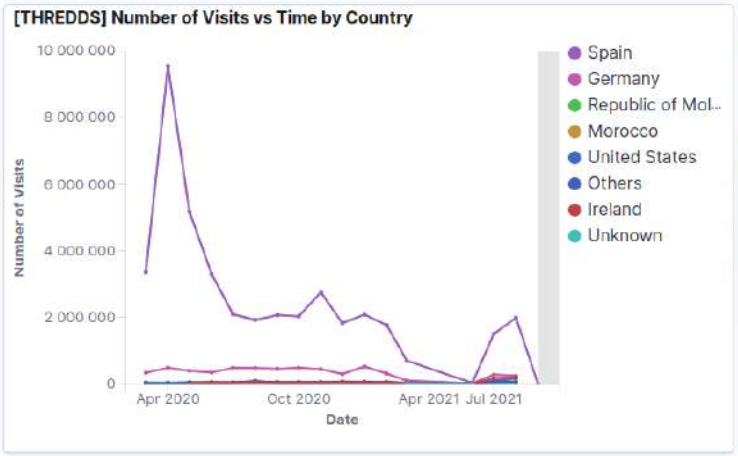
First, the overall access metrics will be presented. Then, traffic by THREDDS protocol, geographic origin, used client, referrer domain and operating system will be analysed.

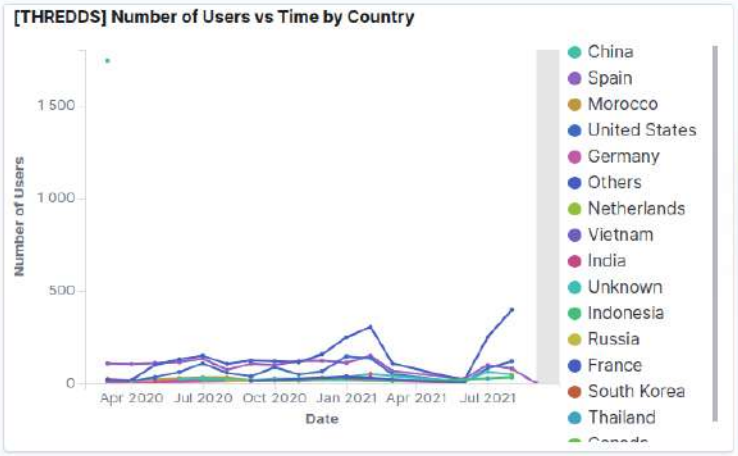
In the next image, the general access metrics in a numeric form are presented.

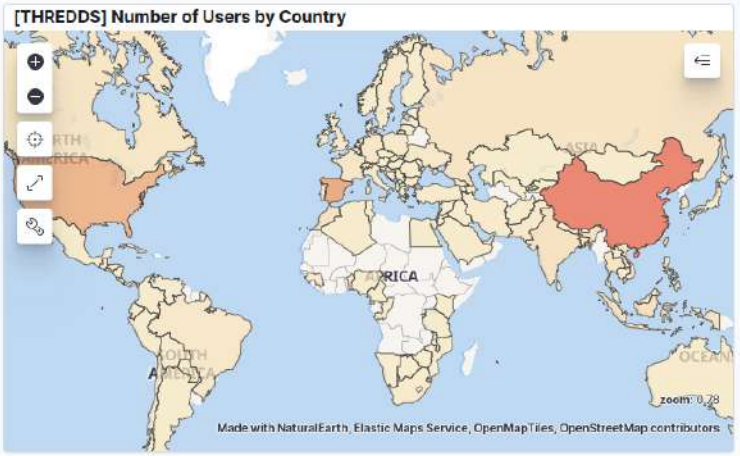


The following image contains the analysis of the traffic by THREDDS protocol (traffic to the catalog HTML pages is assigned to protocol "Catalog" and traffic to the HTML pages of viewers is assigned to "Viewer" protocol).

Next, plots and maps with a geographic analysis of the traffic seen by the server are shown.

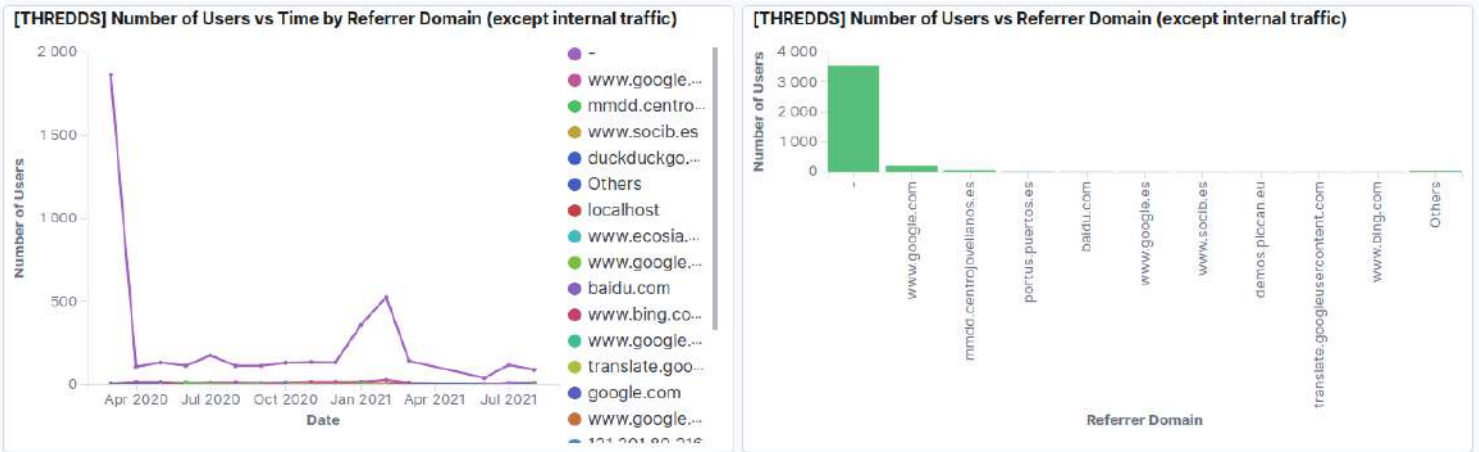




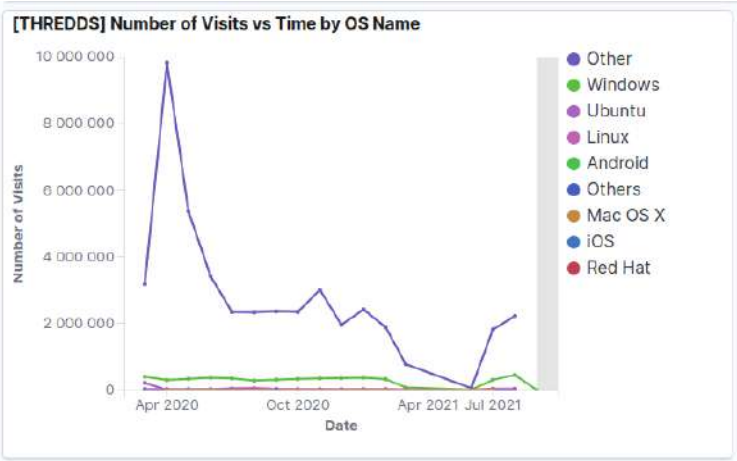


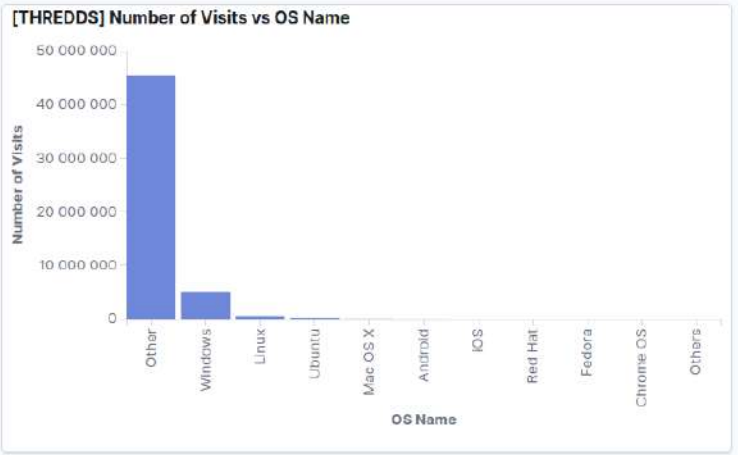
The fourth image shows an analysis of the traffic based on the software used by the user to perform the request (a browser or an HTTP library).

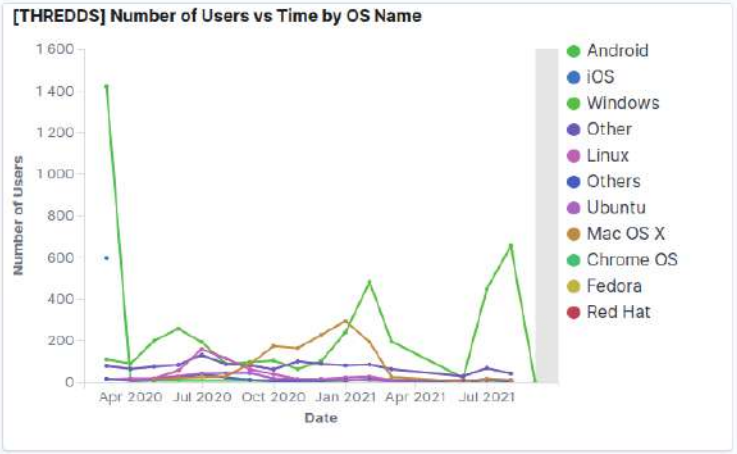
The image below contains an analysis based on the referrer domain ("-" means referrer is not informed, like a visit from a bookmarked link or direct access via software).

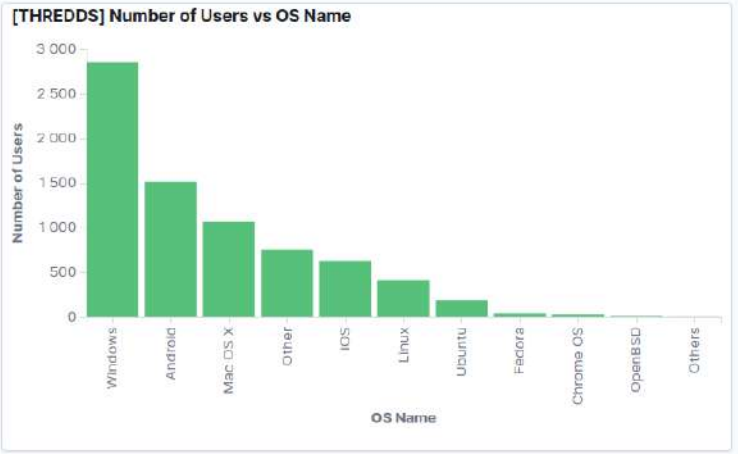


Finally, an analysis of traffic based on the operating system of the user is presented (that information is extracted from the browser version, for HTTP libraries, the information is not available and it is contained in the "Other" category).









Annexe 1.15 Swedish Oceanographic Data Centre (ID 32.1)

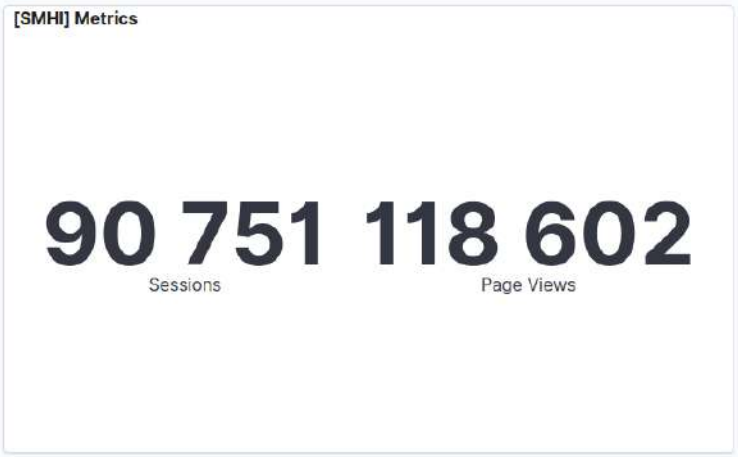
This VA infrastructure consists of two websites: the main SMHI website and OpenData View.

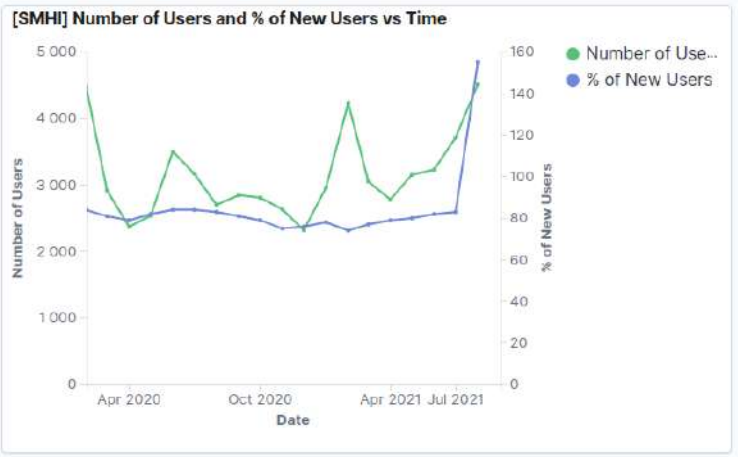
The data underlying the plots of both websites are obtained via a Google Form that SMHI staff kindly fill monthly based on their Google Analytics and Matomo reports.

SMHI's Website

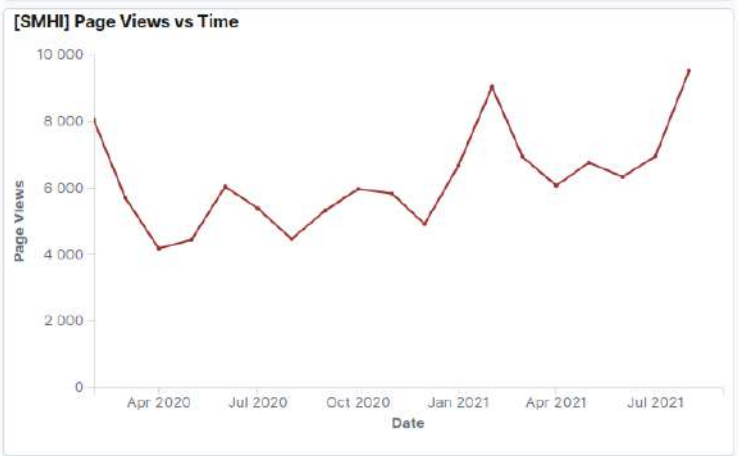
First, the volume of traffic will be analysed. Then, metrics related to session quality (session duration and bounce rate) will be presented. Finally, the origin of traffic will be analysed.

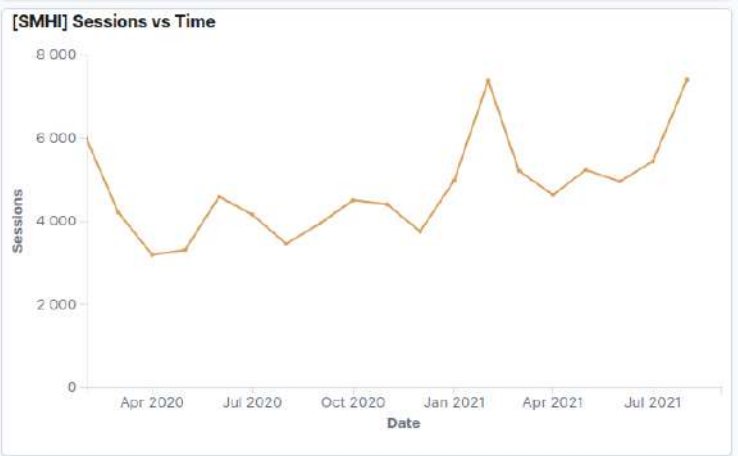
In the next image, a numeric representation of the access metrics and the users analysis is presented.



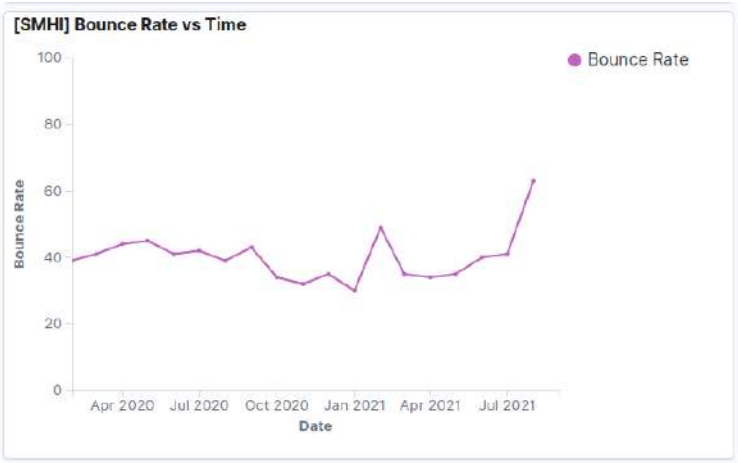


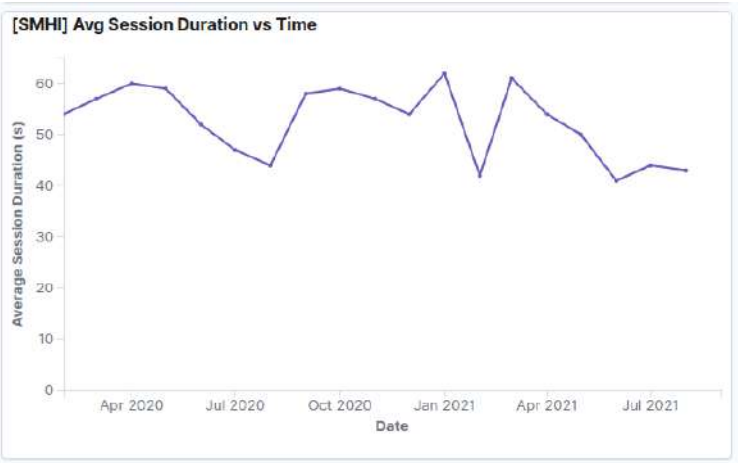
Next, a temporal analysis of the pageviews and sessions has been added.



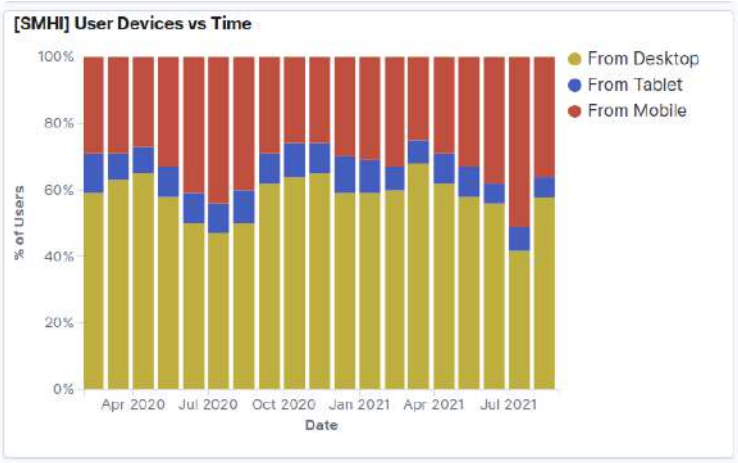


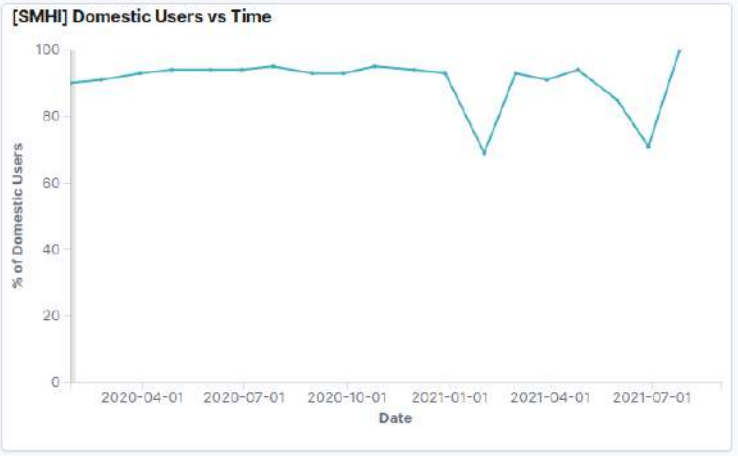
In the following image, metrics related to the on-page user experience: bounce rate and session duration are analysed.





Finally, the user origin is analysed: both by device and country of origin.

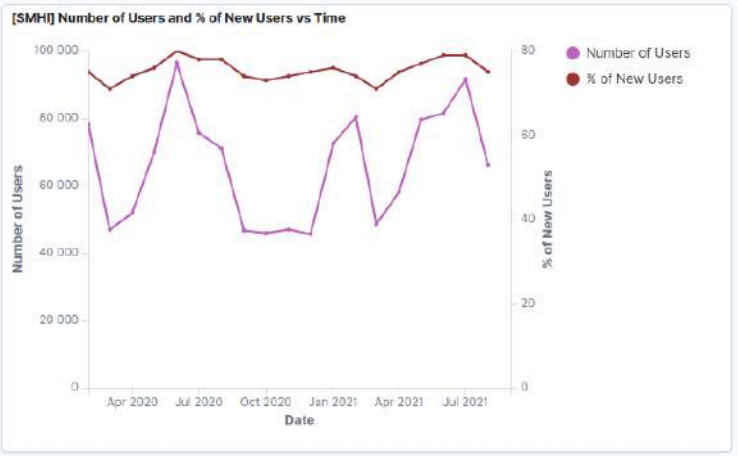




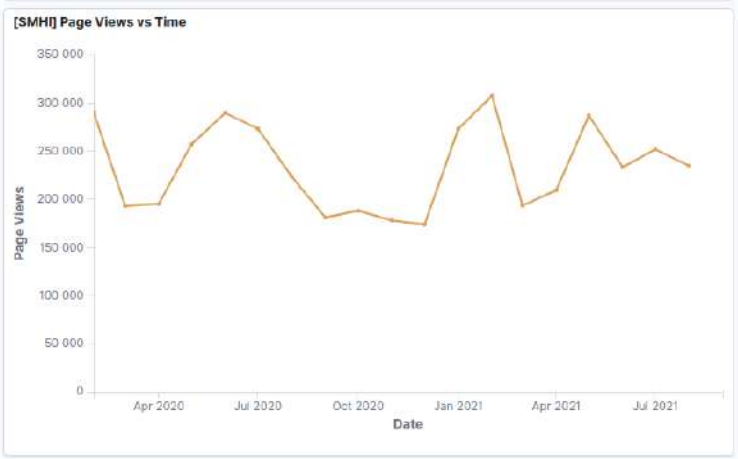
OpenData View Website

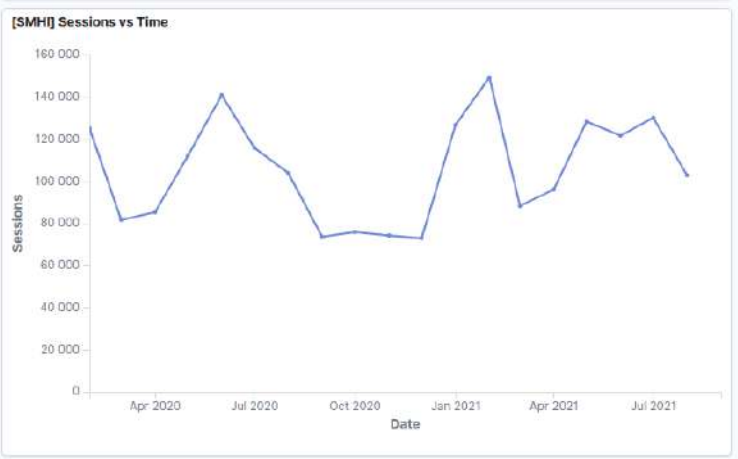
First, the volume of traffic will be analysed. Then, metrics related to session quality (session duration and bounce rate) will be presented. Finally, the origin of traffic will be analysed.

In the next image, a numeric representation of the access metrics and the users analysis is presented.

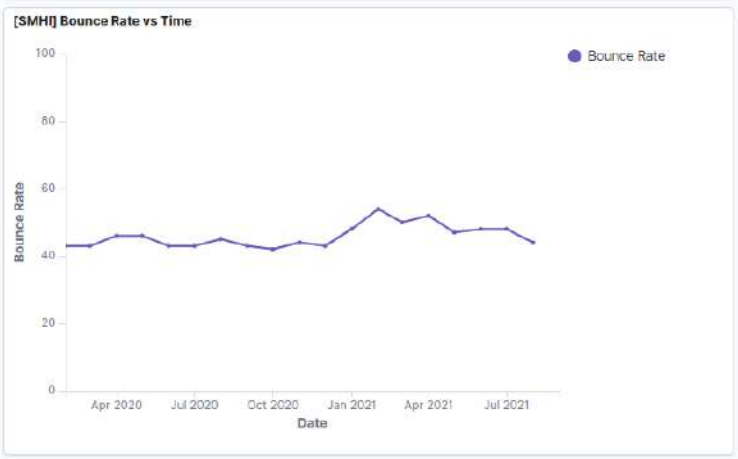


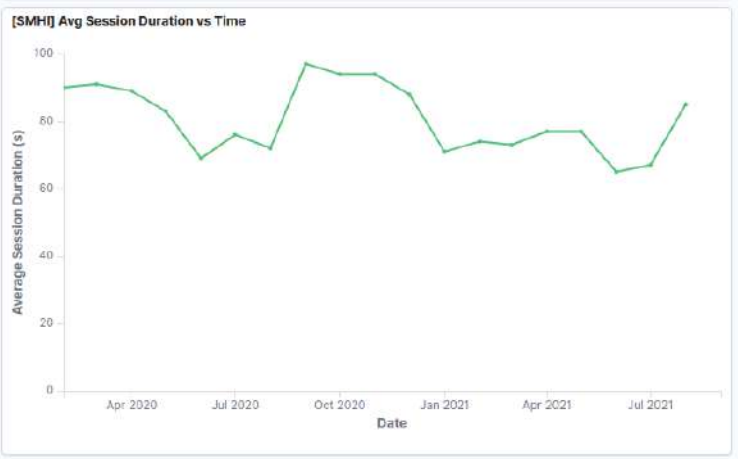
Next, a temporal analysis of the pageviews and sessions has been added.



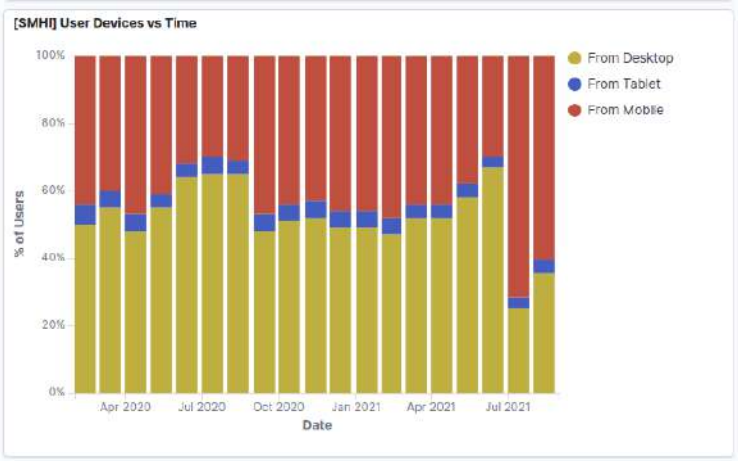


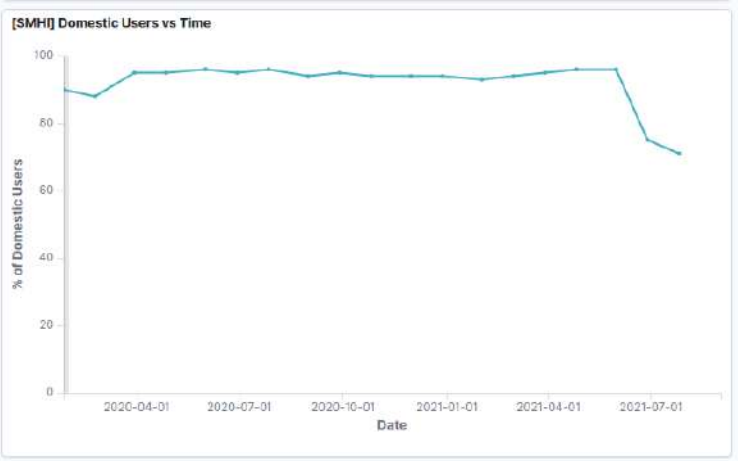
In the following image, metrics related to the on-page user experience: bounce rate and session duration are analysed.





Finally, the user origin is analysed: both by device and country of origin.





Annexe 1.16 SOCIB Data Centre Multi-Platform Observatory (ID 33.1)

This VA infrastructure consists of three parts: a THREDDS data server, a Data Catalog and an API. The THREDDS data server is monitored by analysing its log files and its Google Analytics data. The Data Catalog is monitored using Google Analytics. And the API is monitored by analysing its log files.

THREDDS Data Server

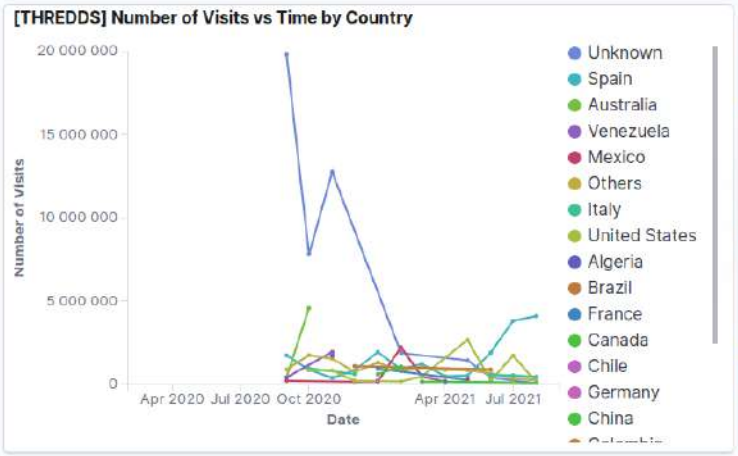
The data underlying the following plots have been obtained from the log files of the THREDDS data server.

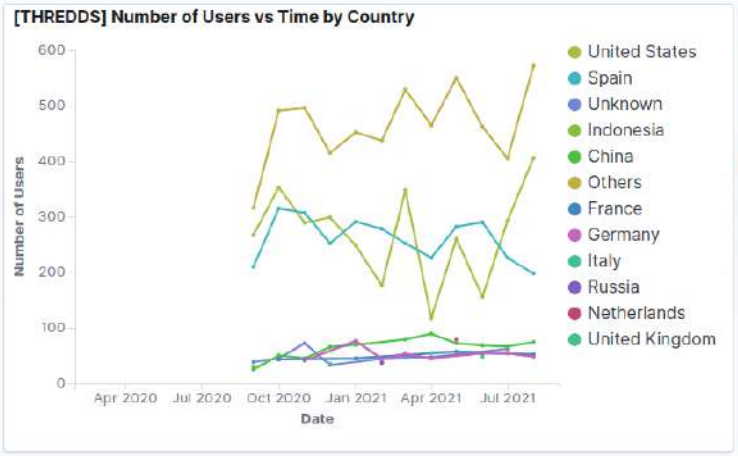
First, the overall access metrics will be presented. Then, traffic by THREDDS protocol, geographic origin, used client, referrer domain and operating system will be analysed. Thanks to the fact that URLs are formatted in a known way, additional categories can be added to each visit. This is done by parsing the URL and extracting additional information. Therefore, the next images present the analysis of pageviews and users over time and overall by: Observing System, Platform Type, Platform Name, Instrument Name, Processing Level and Aggregation Level.

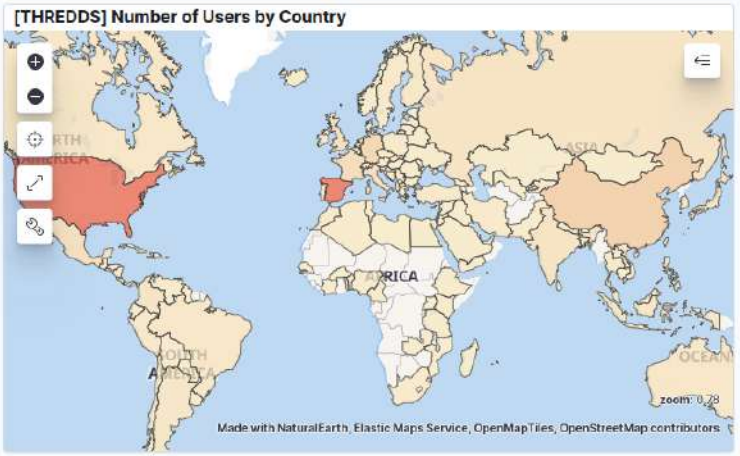
In the next image, the general access metrics in a numeric form are presented.

The following image contains the analysis of the traffic by THREDDS protocol (traffic to the catalog HTML pages is assigned to protocol "Catalog" and traffic to the HTML pages of viewers is assigned to "Viewer" protocol).

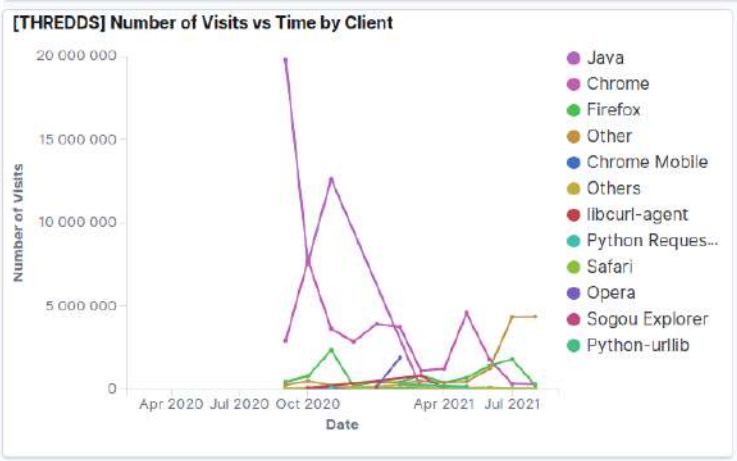
Next, plots and maps with a geographic analysis of the traffic seen by the server are shown.

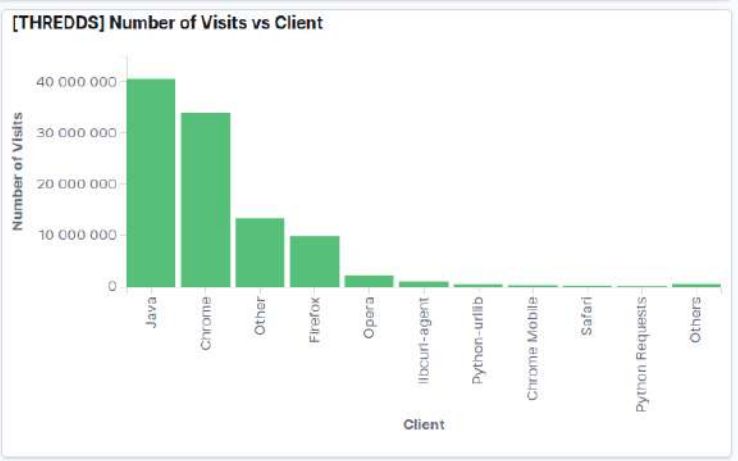


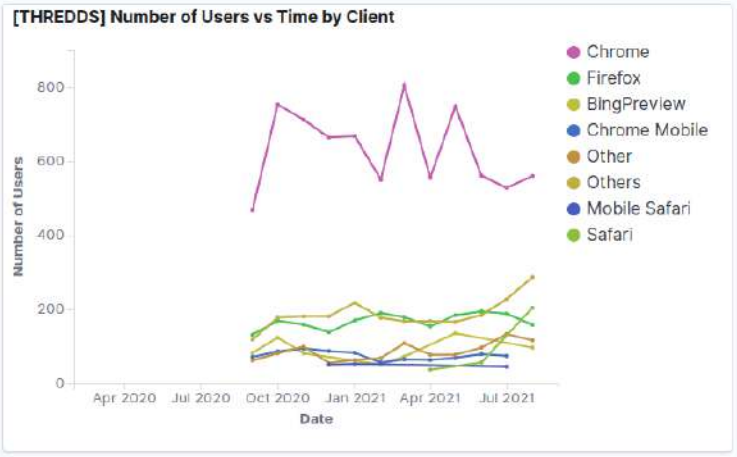


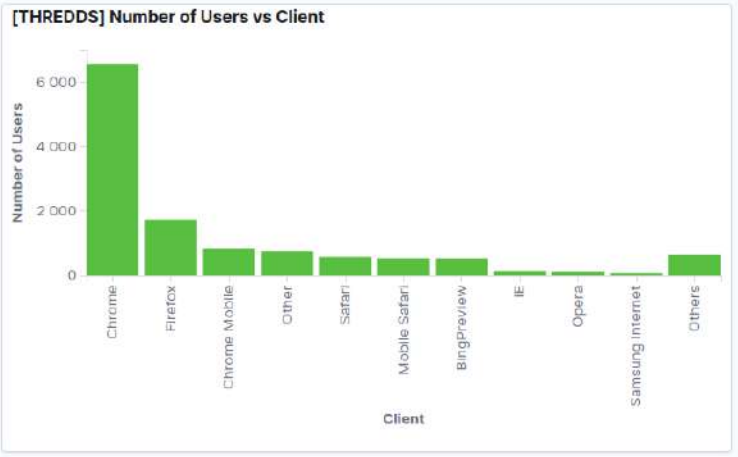


The fourth image shows an analysis of the traffic based on the software used by the user to perform the request (a browser or an HTTP library).

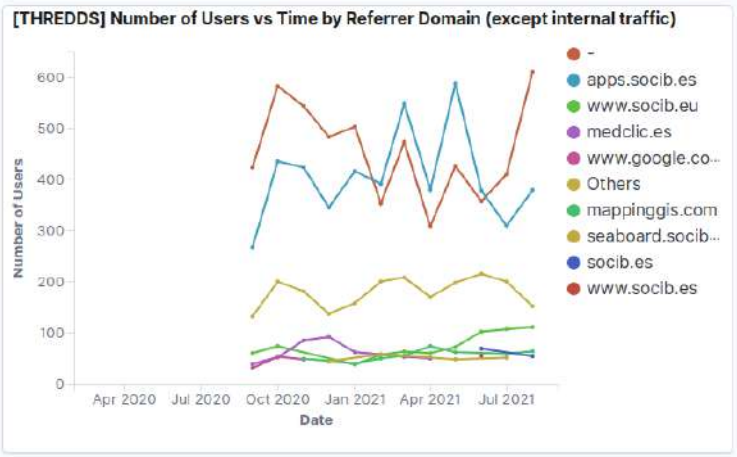


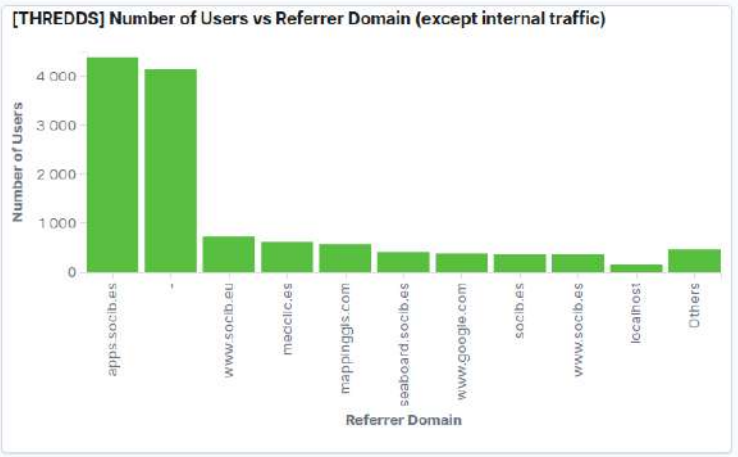




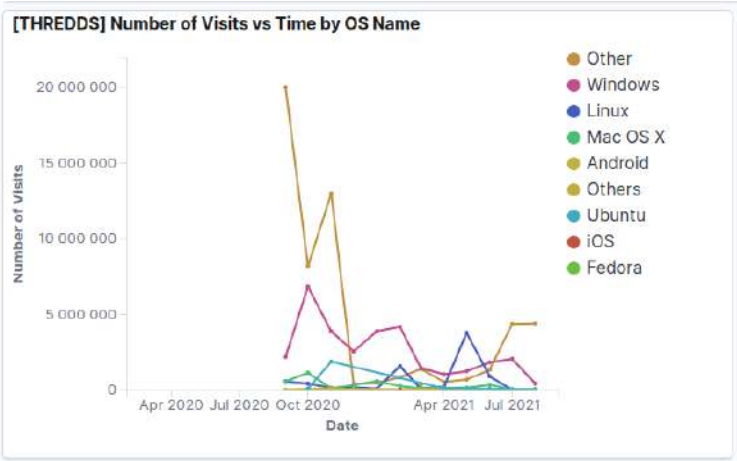


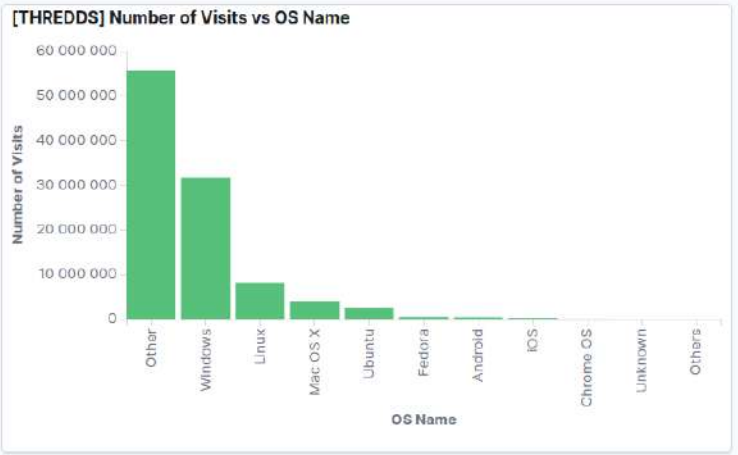
The image below contains an analysis based on the referrer domain ("-" means referrer is not informed, like a visit from a bookmarked link or direct access via software).

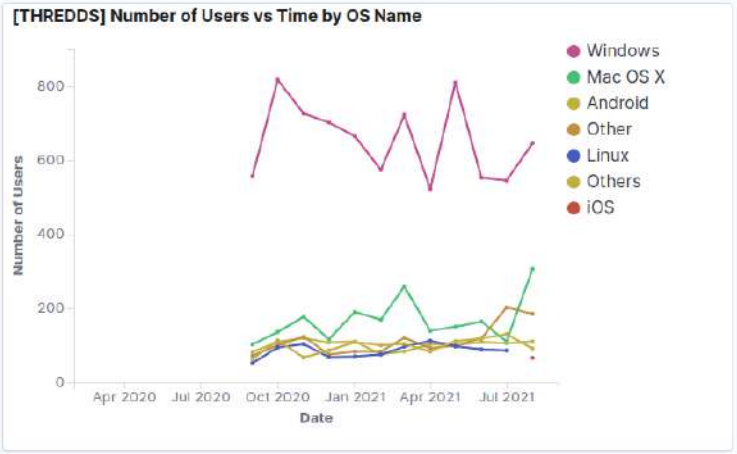


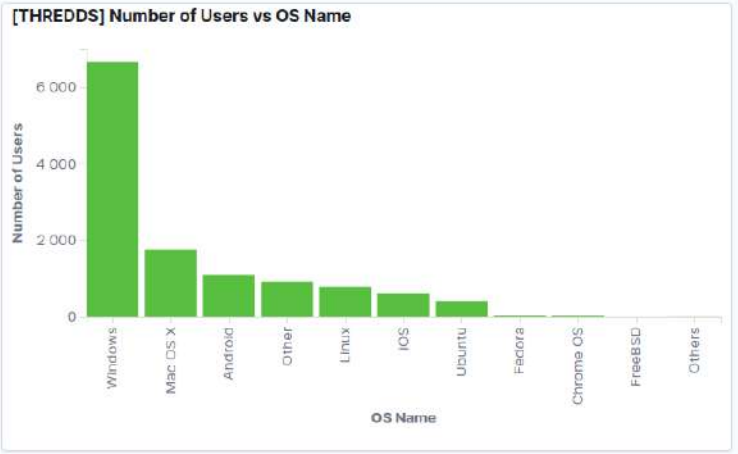


Next, an analysis of traffic based on the operating system of the user is presented (that information is extracted from the browser version, for HTTP libraries, the information is not available and it is contained in the "Other" category).

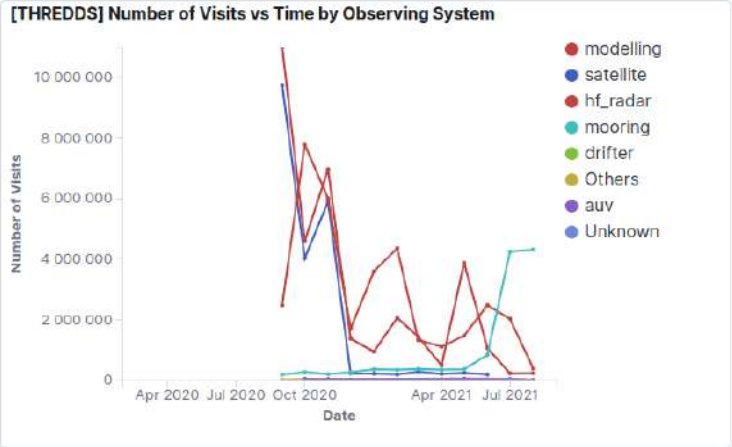


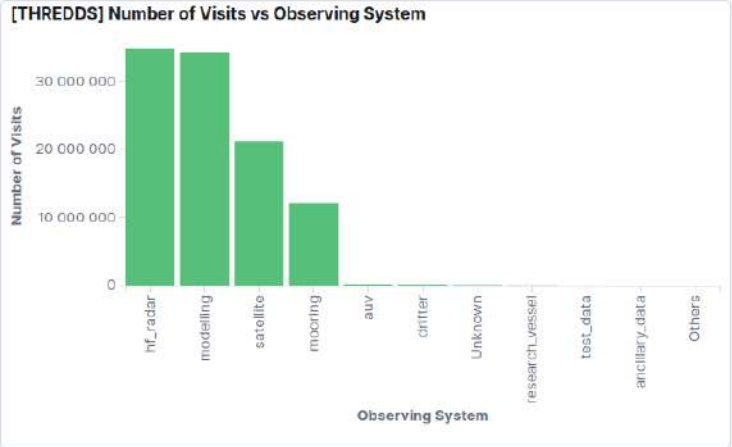


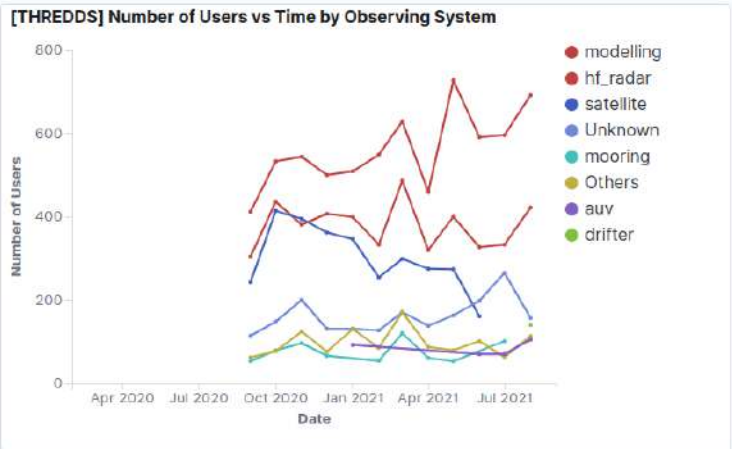


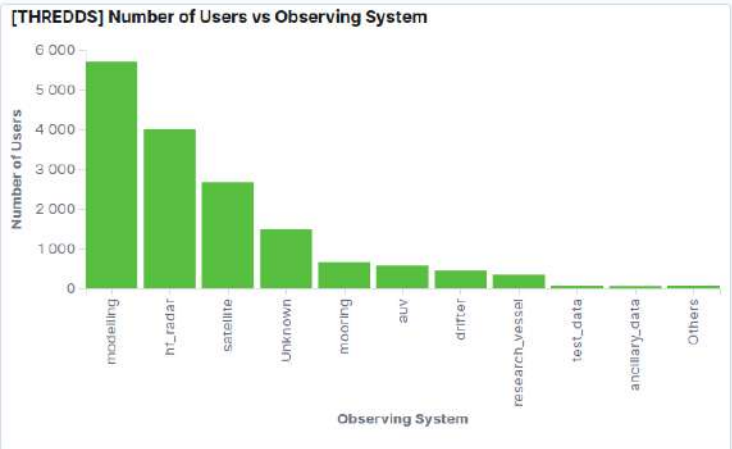


The following is an analysis of the traffic based on the observing system of SOCIB.



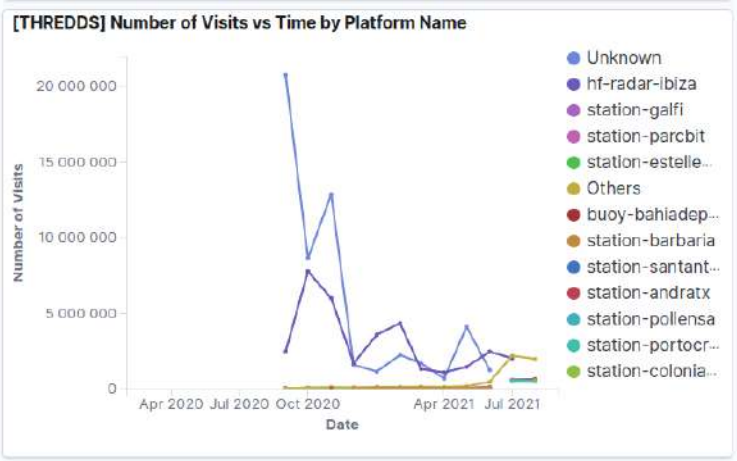


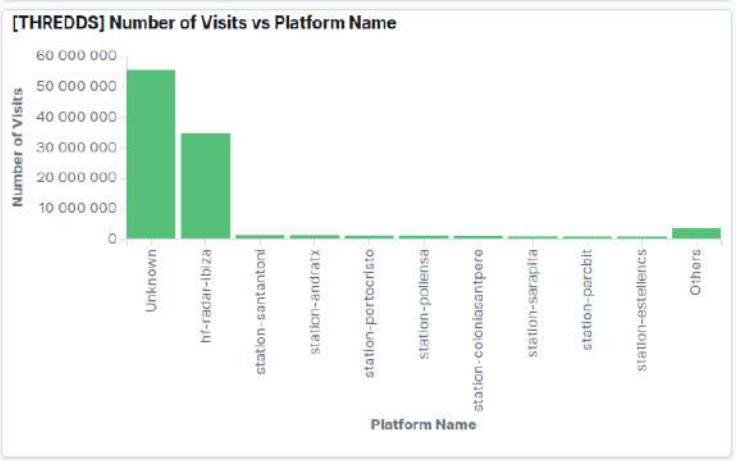


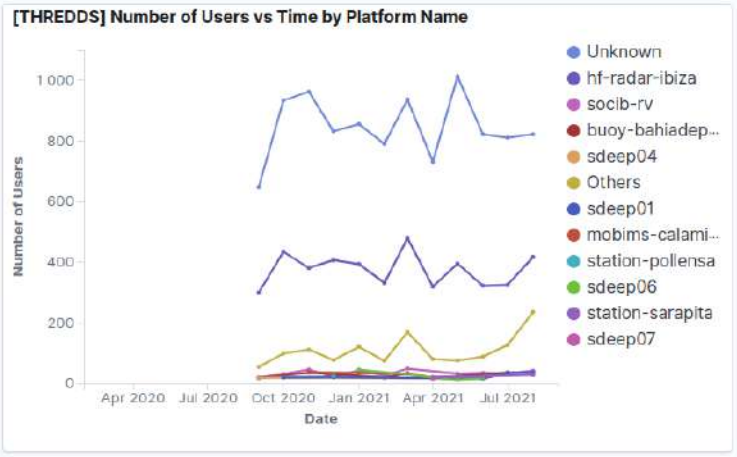


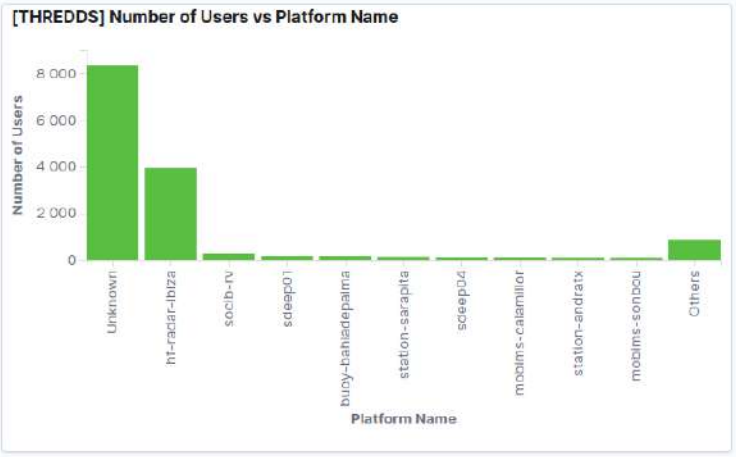
The image below contains an analysis of the traffic based on the type of platform that collected the served data.

The next image presents an analysis of the traffic based on the name of the platform that collected the served data.

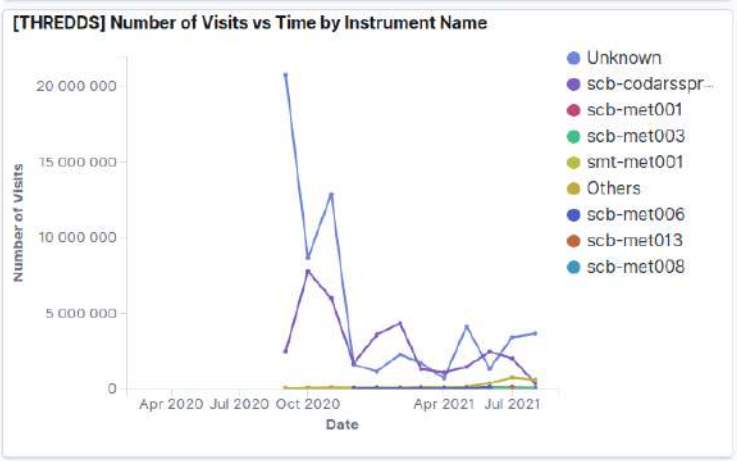


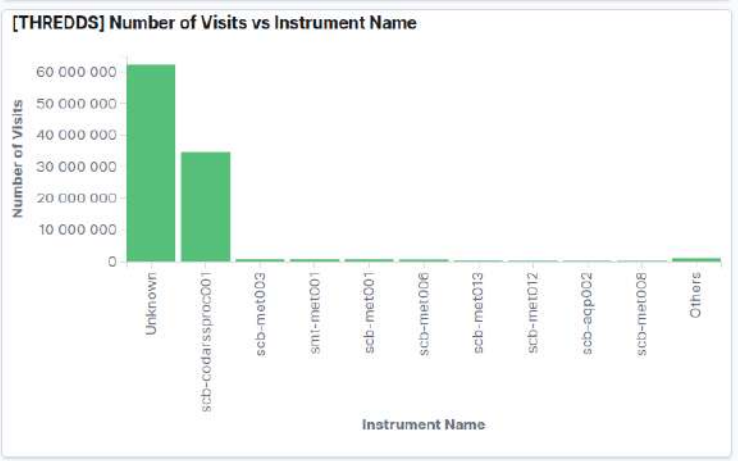


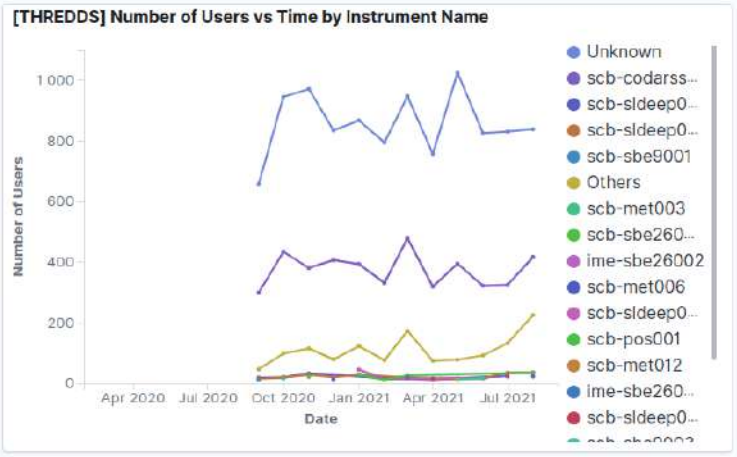


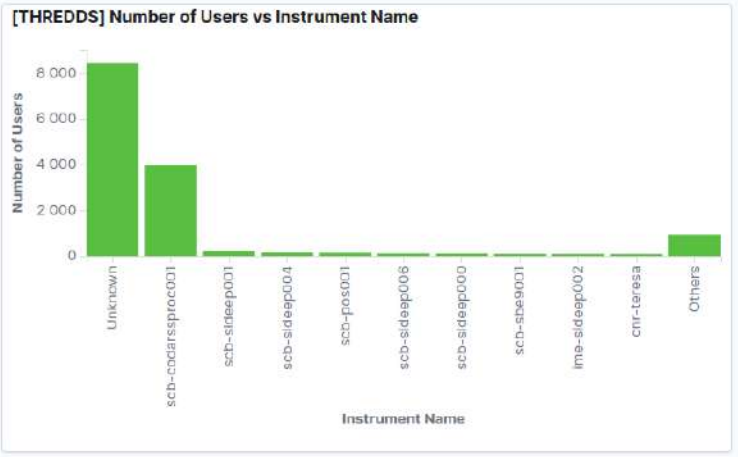


Next, an analysis of the traffic based on the name of the instrument that collected the served data is presented.

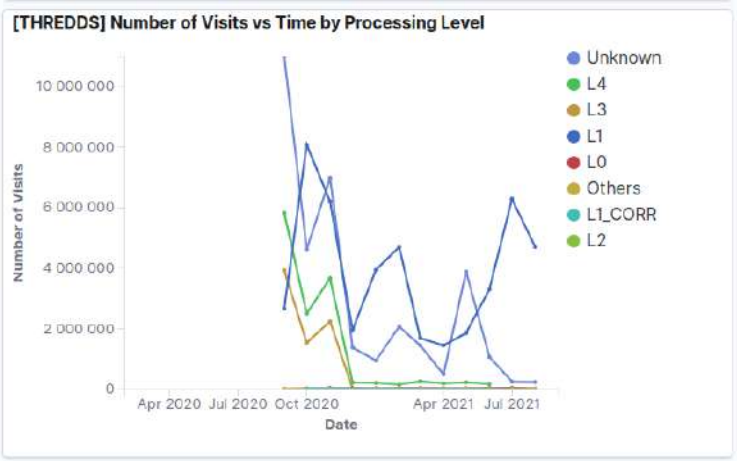


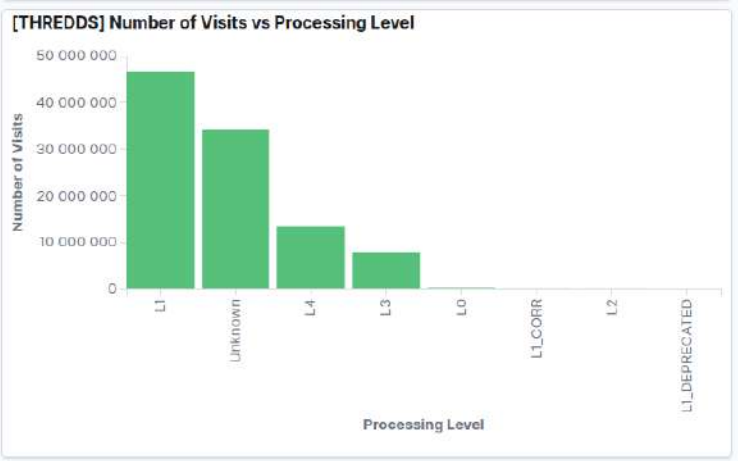


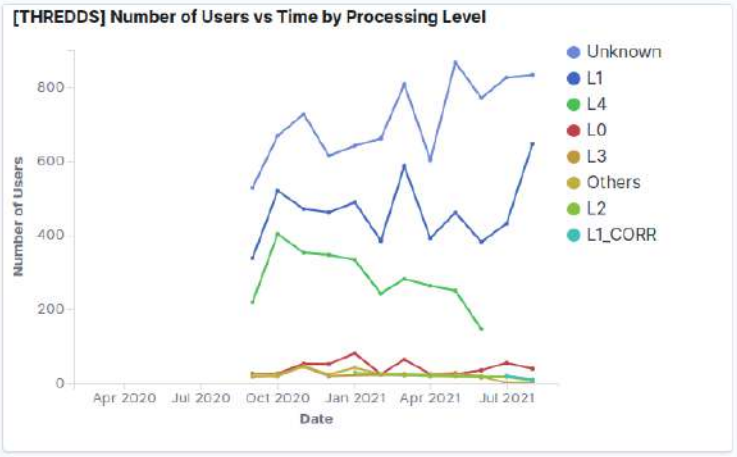


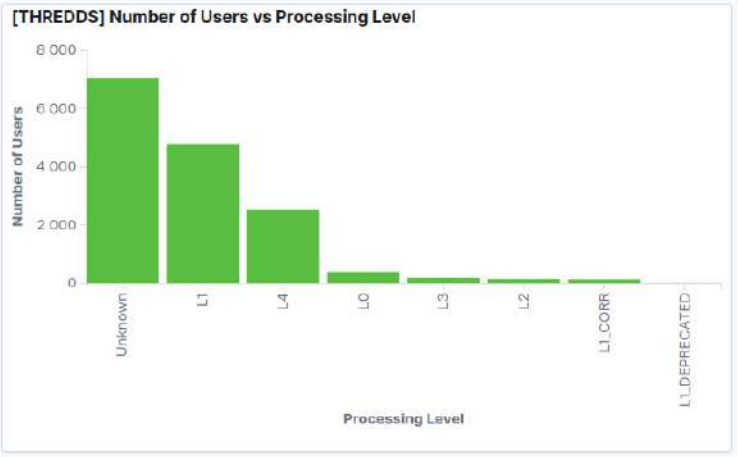


Below, an analysis of the traffic based on the processing level of the served data is presented.

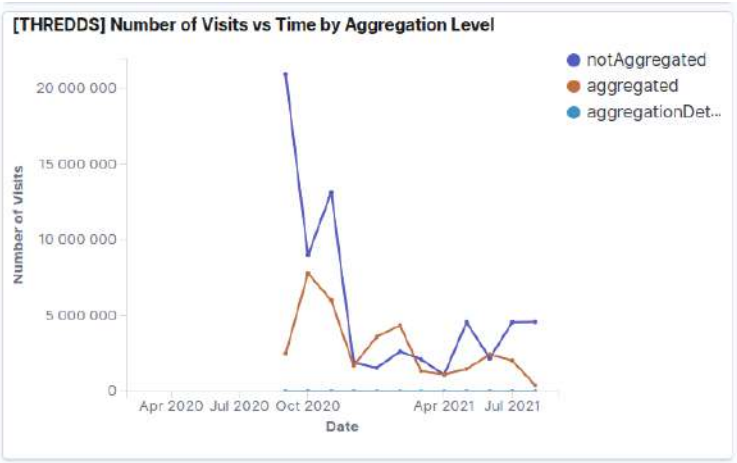


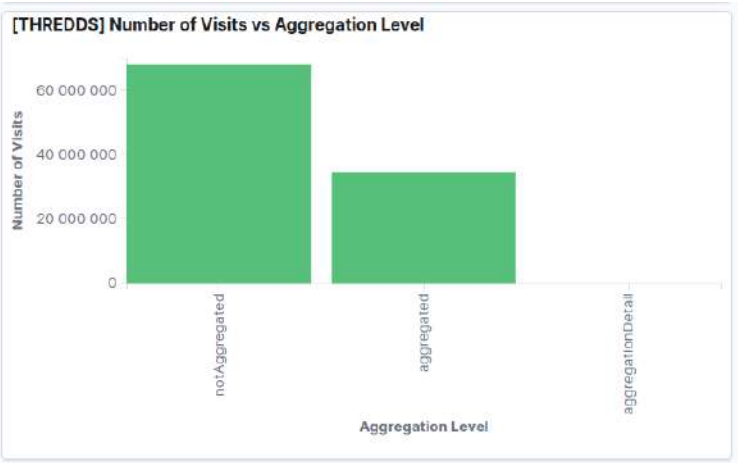


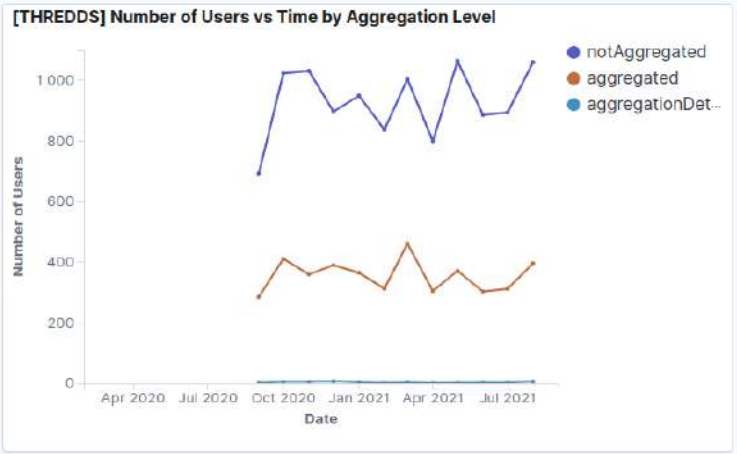


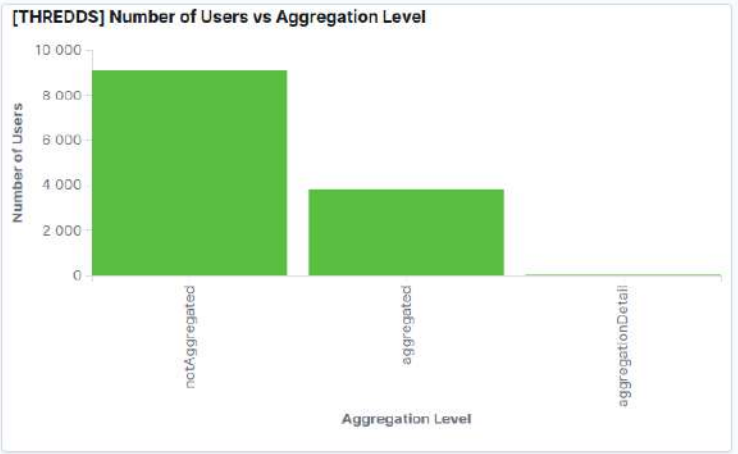


Finally, an analysis of the traffic based on the aggregation level of the served data is presented.







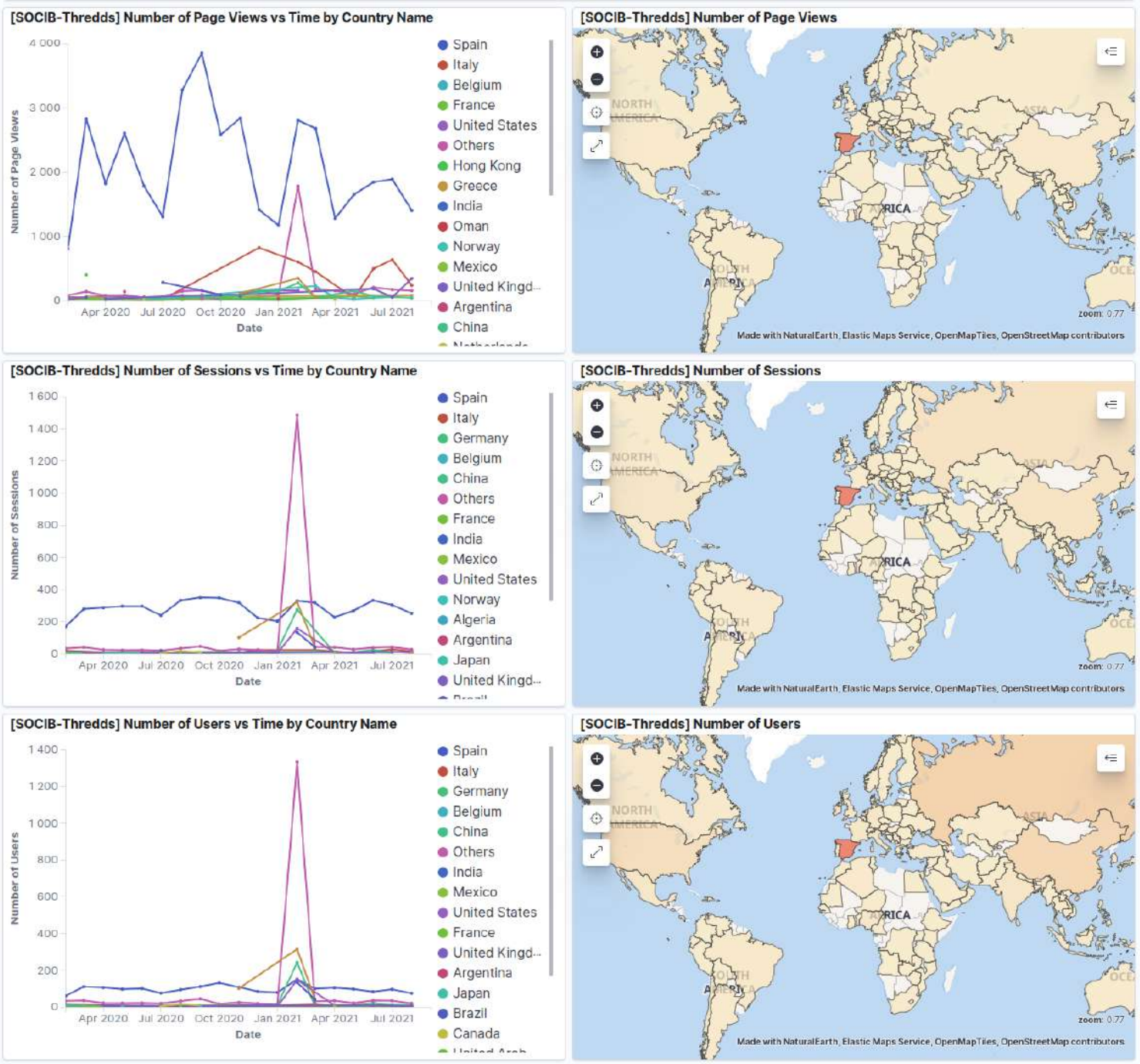


SOCIB's THREDDS data server is also monitored via Google Analytics. This tool gives a different perspective as it only detects the traffic of browsers (or other clients that execute Javascript) and it provides information that cannot be obtained from logs like the Sessions analysis

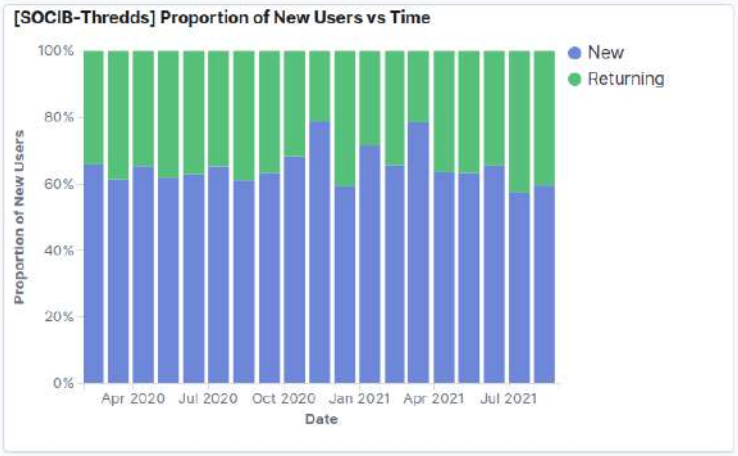
First, the overall access metrics will be presented. Then, the following analysis are included: geographic analysis, analysis by user type (new vs returning), bounce rate analysis, user device analysis and traffic source analysis.

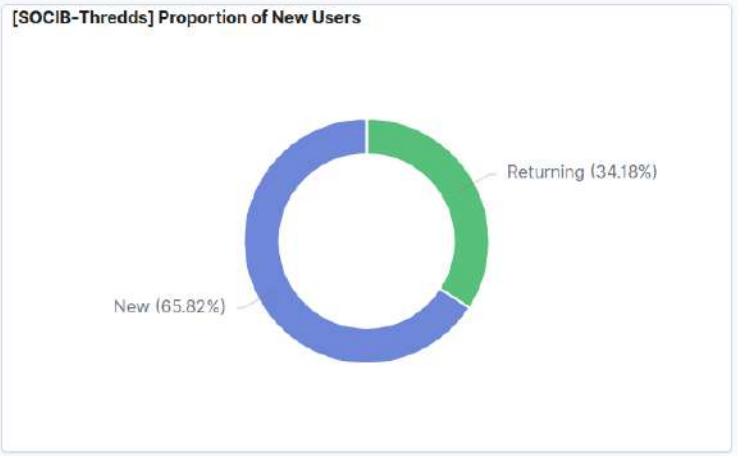
In the next image, the general access metrics in a numeric form are presented.

The following image contains the geographic analysis of the traffic.

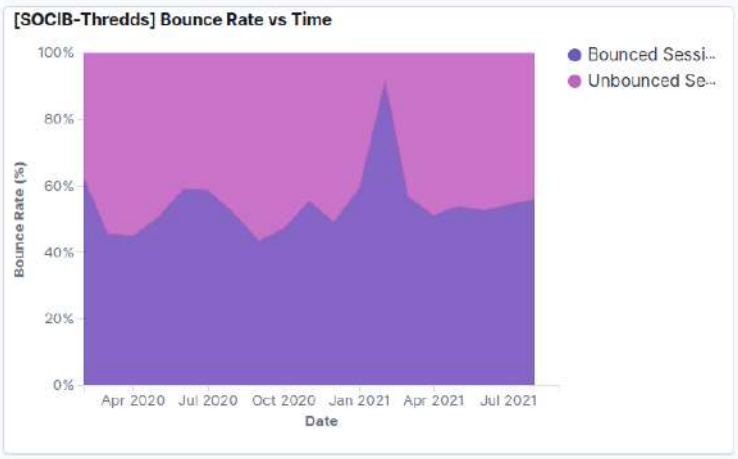


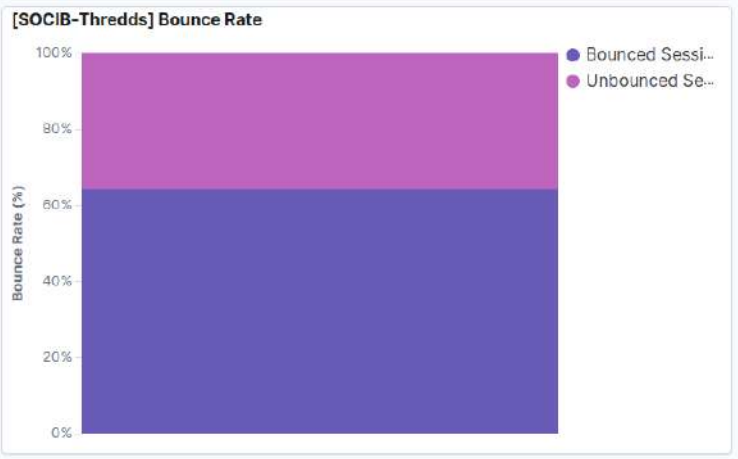
Next, plots analysing the proportion of new users are shown.





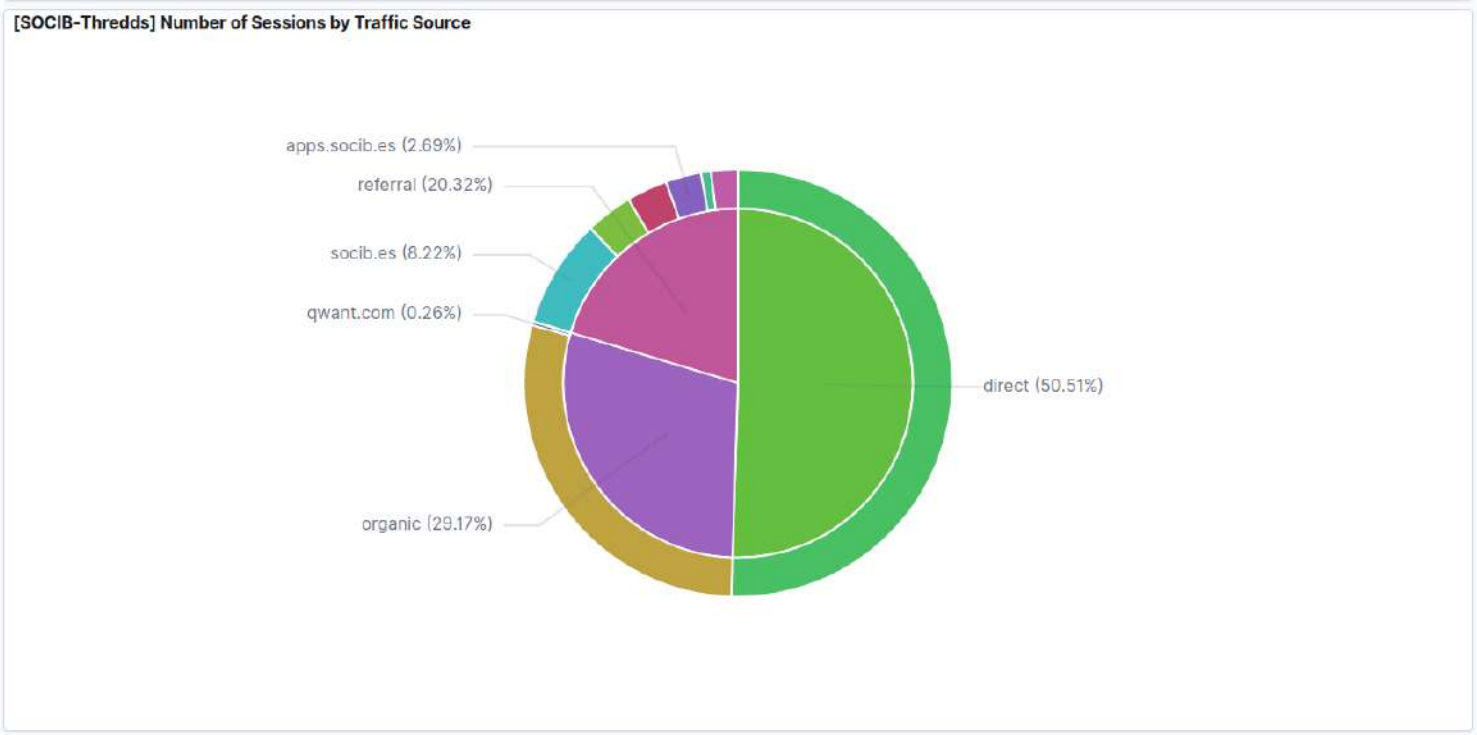
Next, the bounce rate is analysed.





The image below shows the proportion of sessions that come from each type of device and the influence of the type of device on the bounce rate.

Finally, an analysis of traffic based on the source is presented.



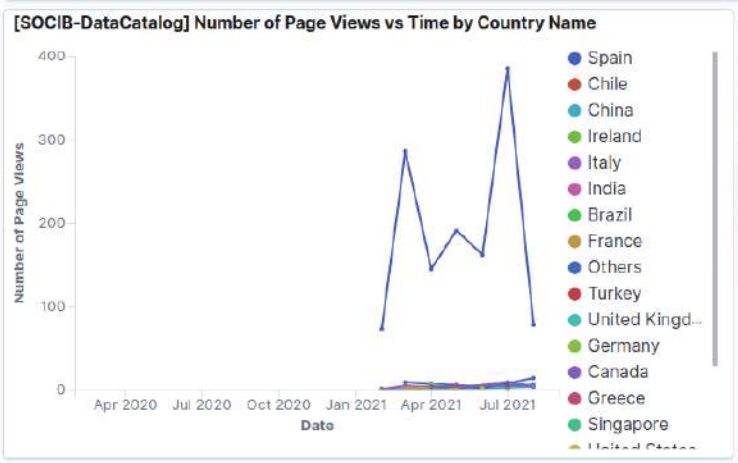
Data Catalog

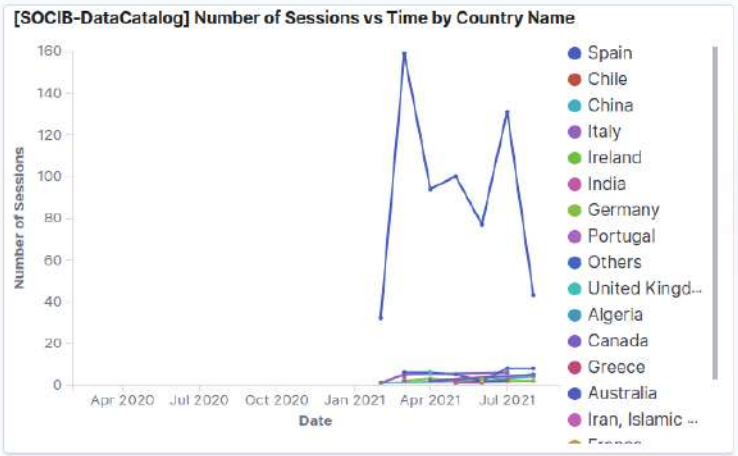
The data underlying the following plots have been obtained from the Google Analytics API.

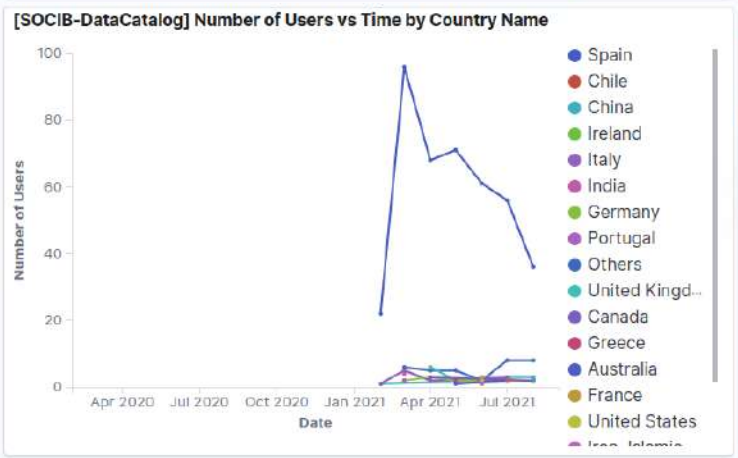
First, the overall access metrics will be presented. Then, the following analysis are included: geographic analysis, analysis by user type (new vs returning), bounce rate analysis, user device analysis and traffic source analysis.

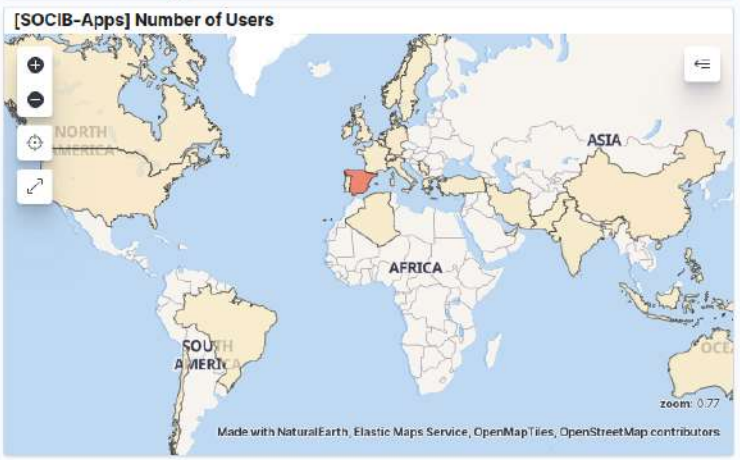
In the next image, the general access metrics in a numeric form are presented.

The following image contains the geographic analysis of the traffic.

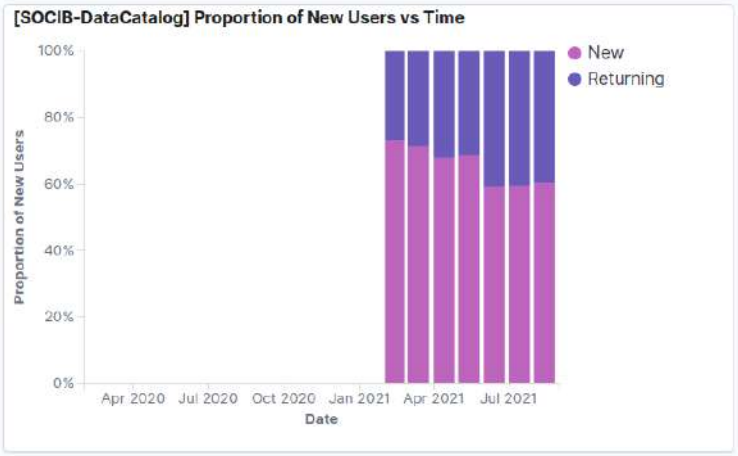


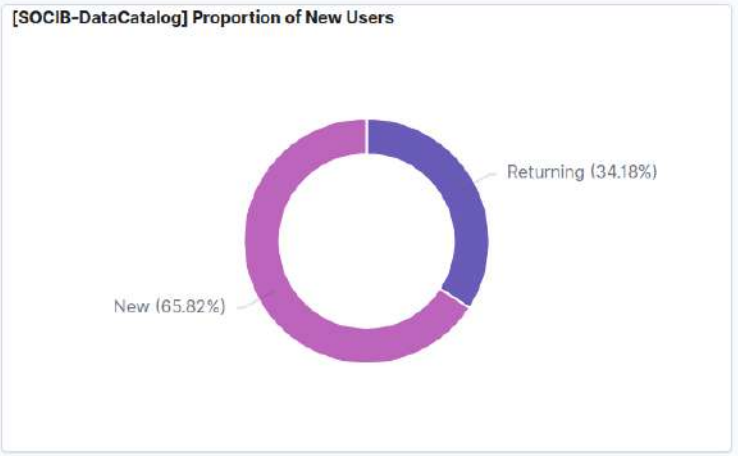




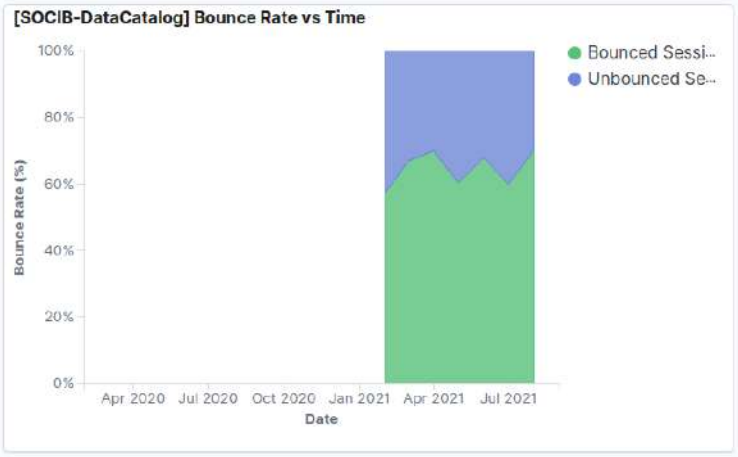


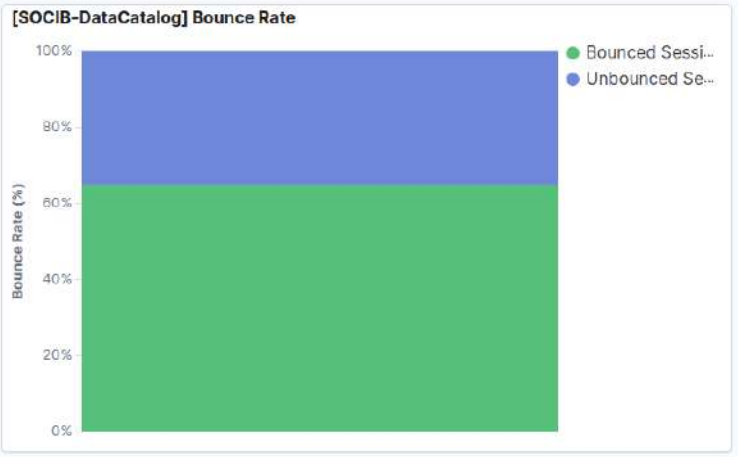
Next, plots analysing the proportion of new users are shown.



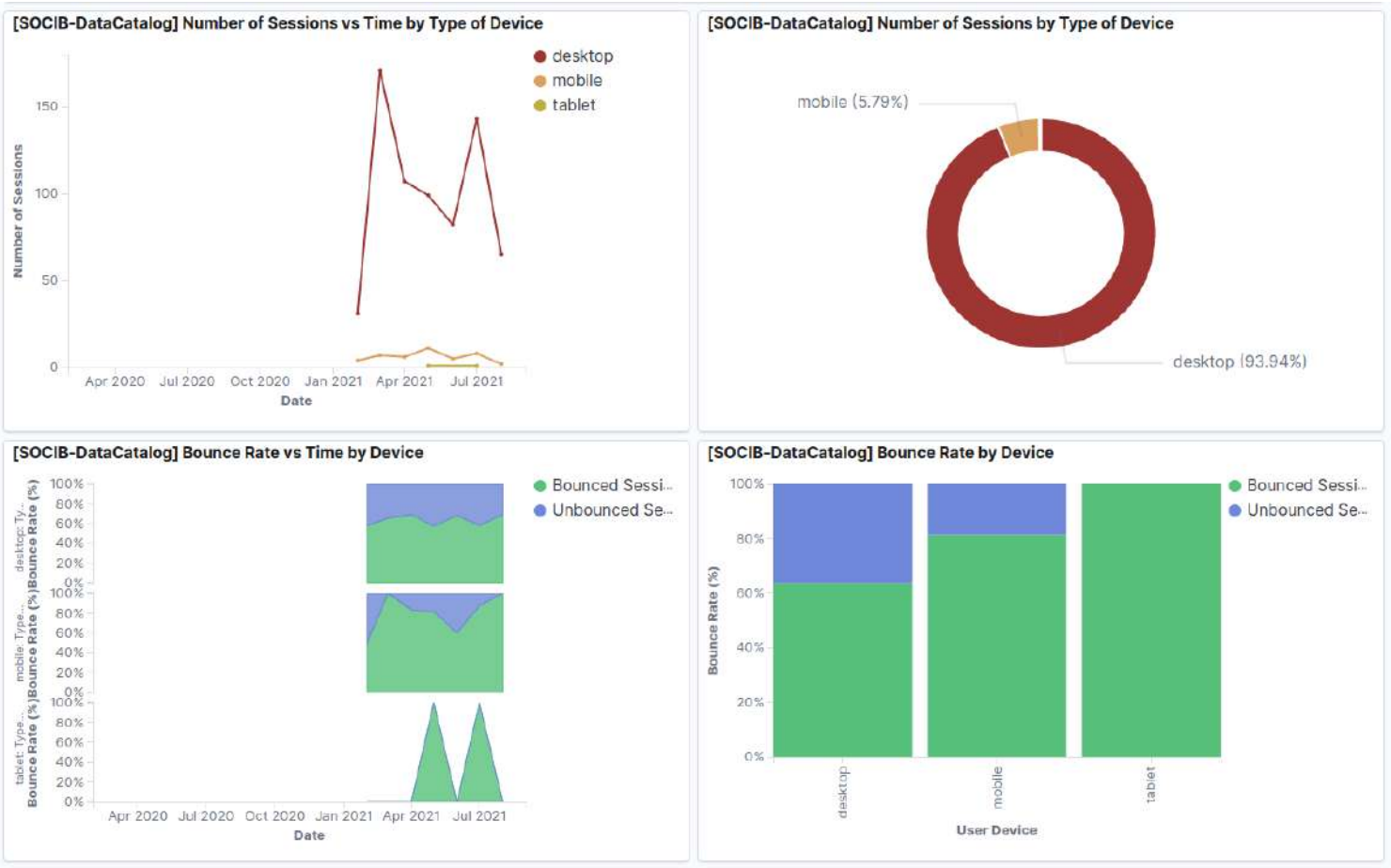


Next, the bounce rate is analysed.

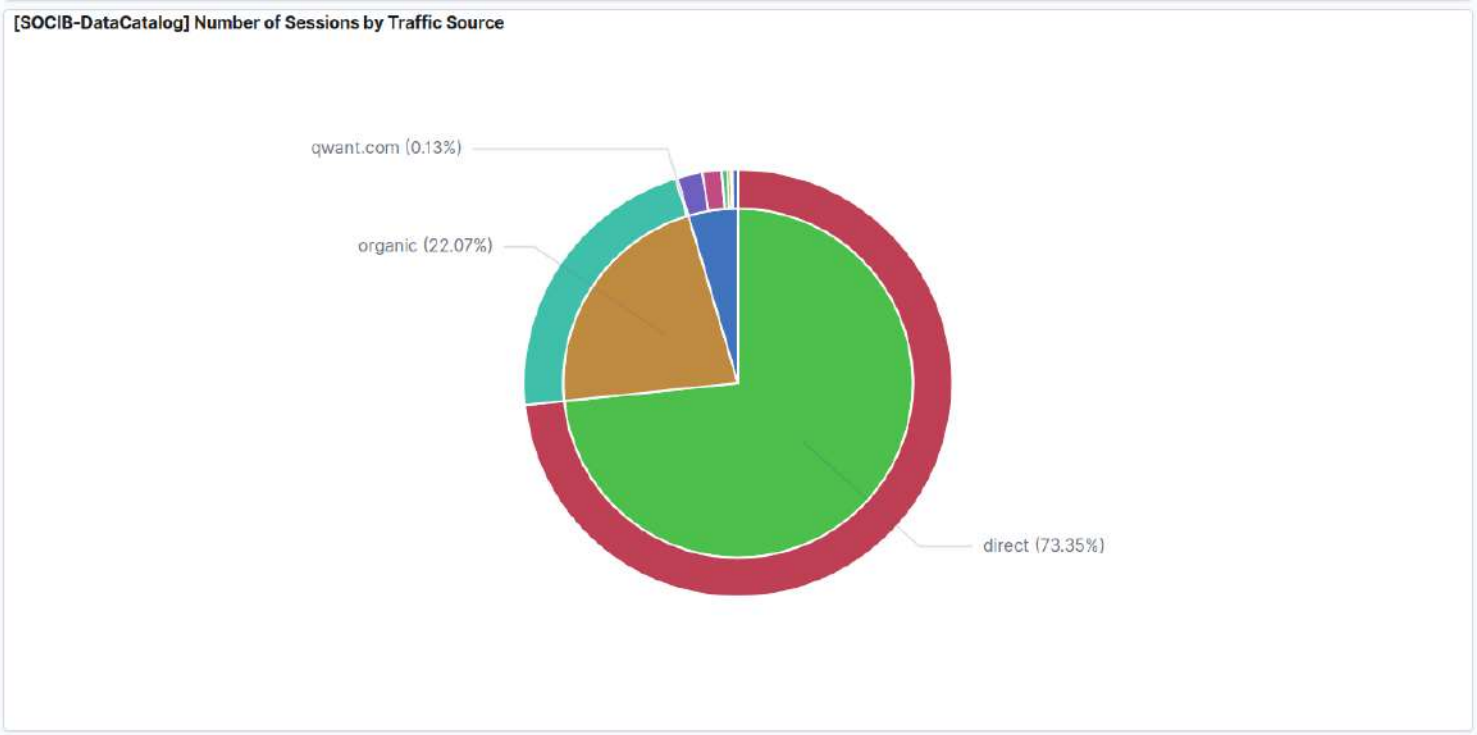




The image below shows the proportion of sessions that come from each type of device and the influence of the type of device on the bounce rate.



Finally, an analysis of traffic based on the source is presented.

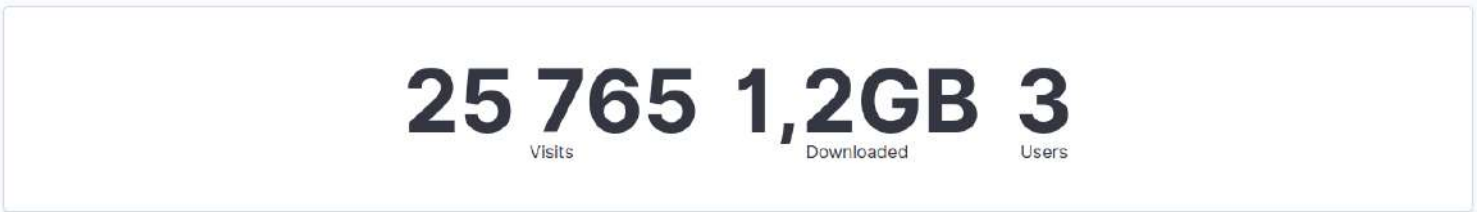


API

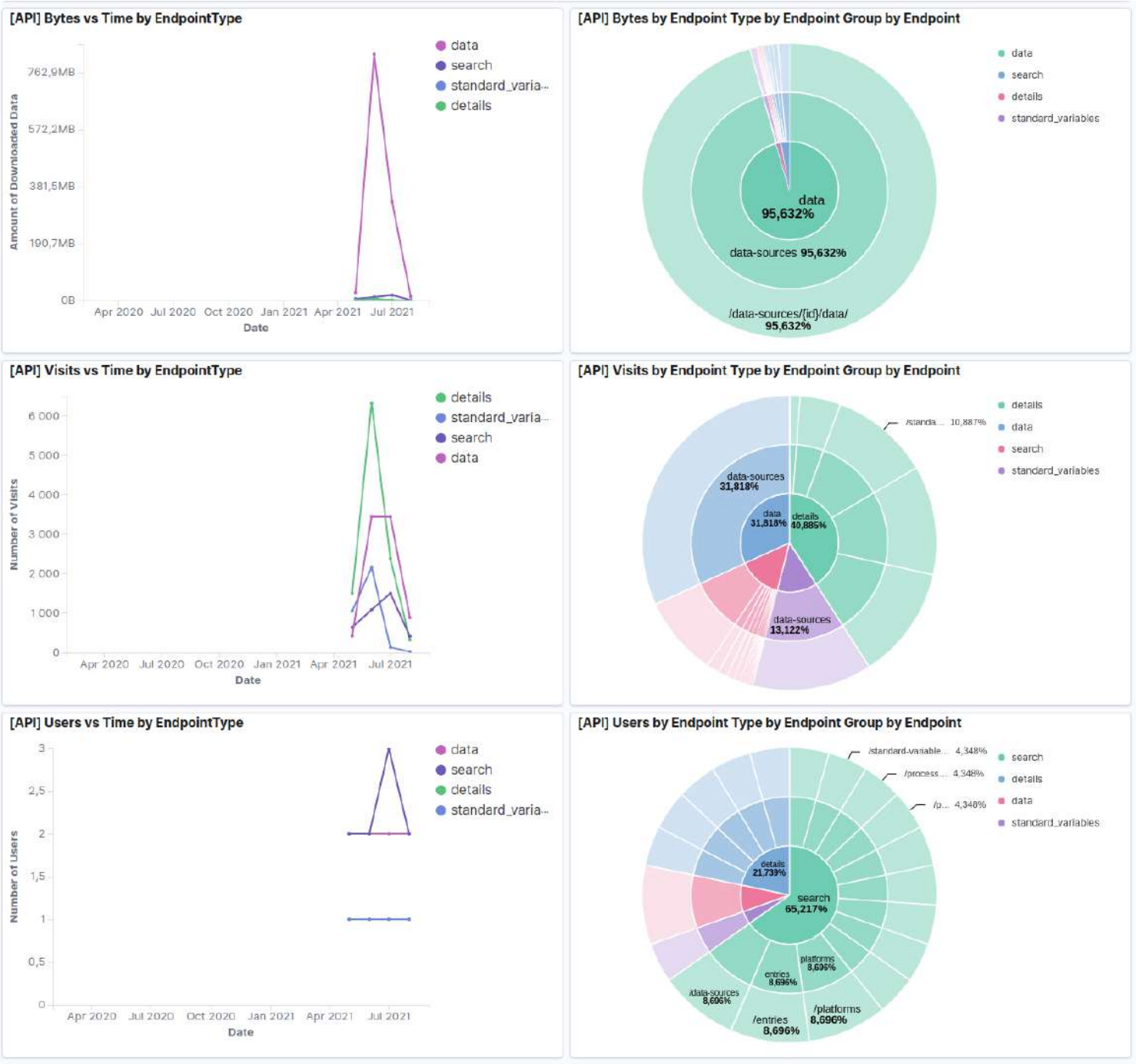
The data underlying the following plots have been obtained from the log files of the API server. Since the API requires authentication, distinct IPs are not used to estimate the number of users, the actual number of users that access the API is used. This API was launched on May 10th 2021, hence the unavailable data at the beginning of JERICO-S3. This API is, essentially, in the testing phase and its access metrics are included here for completeness. No conclusions can be taken from this small amount of data.

First, the overall access metrics will be presented. Then, traffic will be analysed by content, geographic origin, used client, referrer domain and operating system.

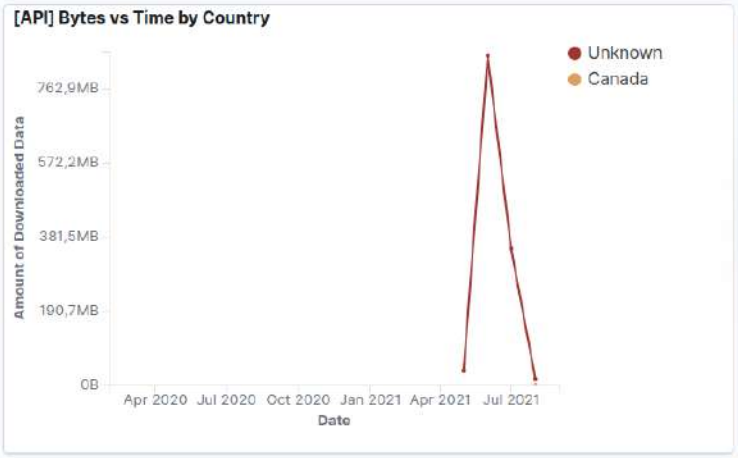
In the next image, the general access metrics in a numeric form are presented.

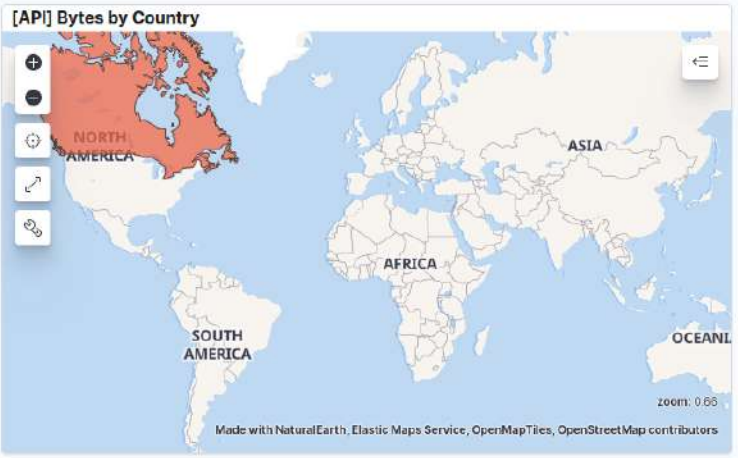


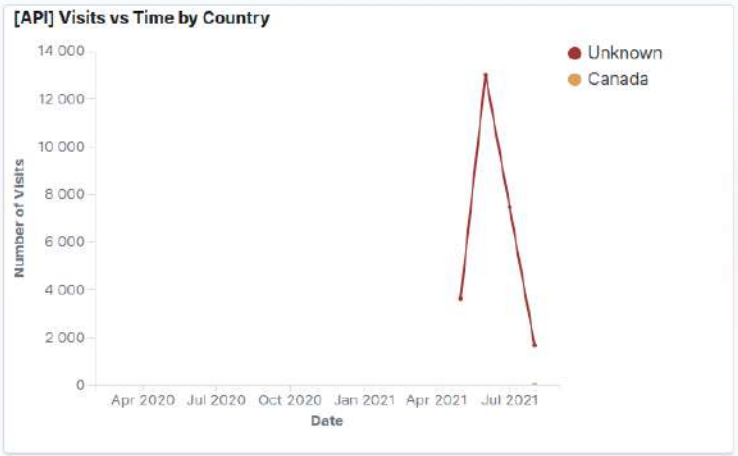
In the following image, the content analysis is presented. It can be seen what proportion of data, pageviews and users goes to each family of endpoints (data, search, details and standard_variables). It can also be seen how the traffic to API endpoints is distributed among the types of endpoints and what specific endpoints are being visited.

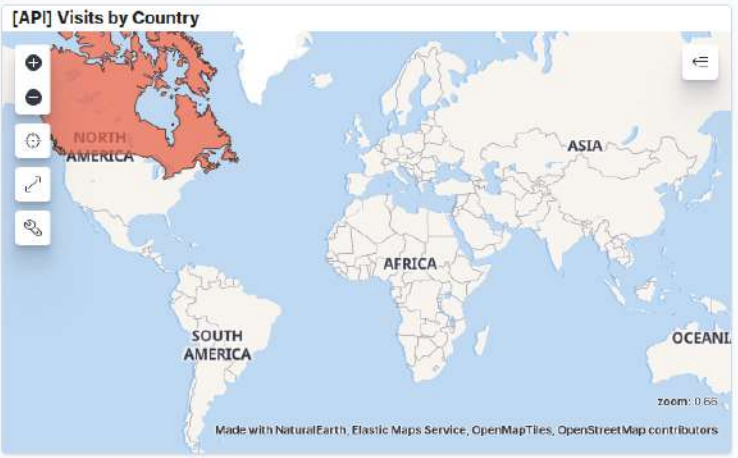


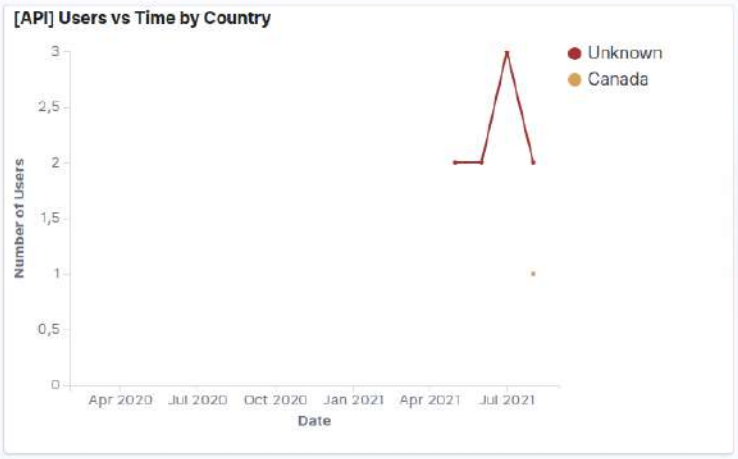
Next, plots and maps with a geographic analysis of the traffic seen by the server are shown.

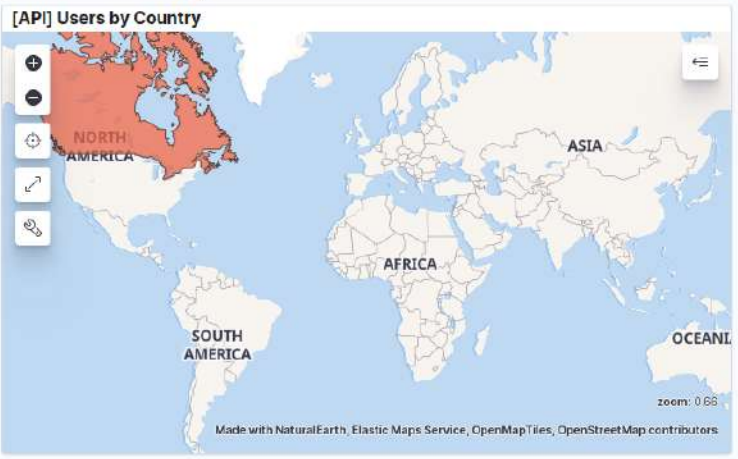




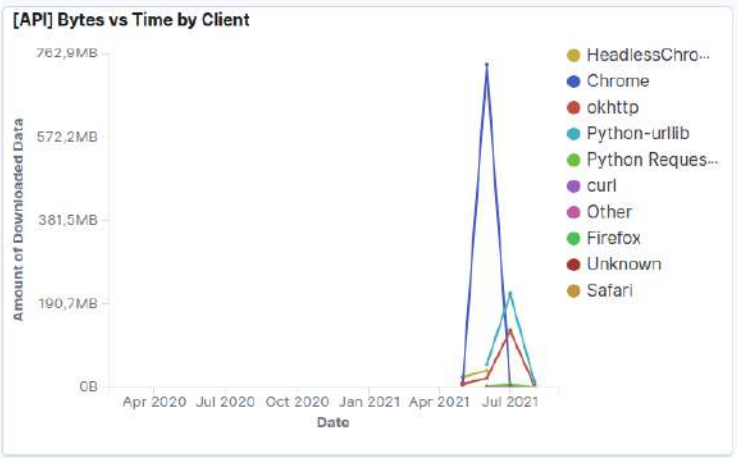


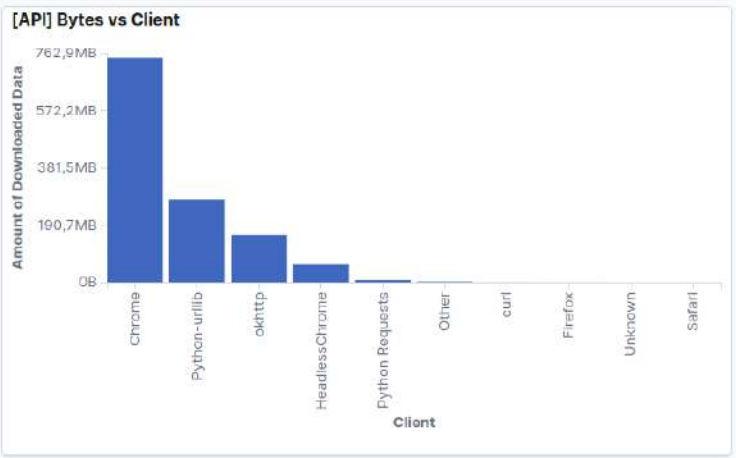


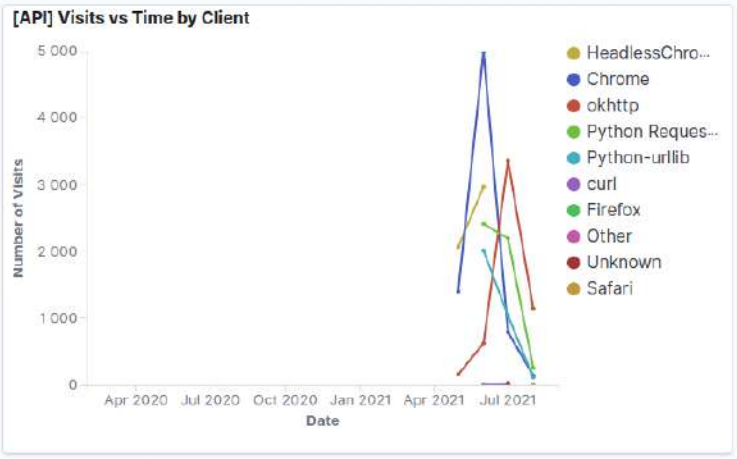


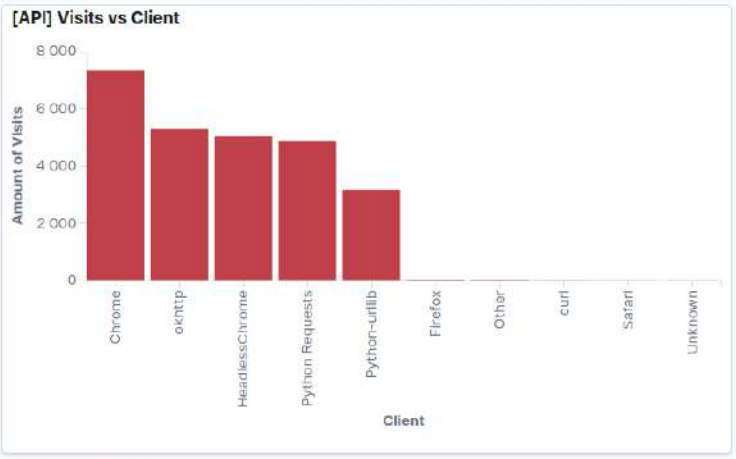


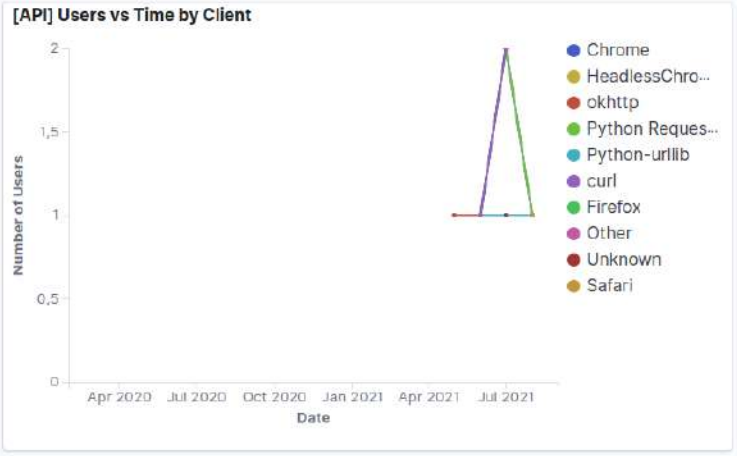
The fourth image shows an analysis of the traffic based on the software used by the user to perform the request (a browser or an HTTP library).

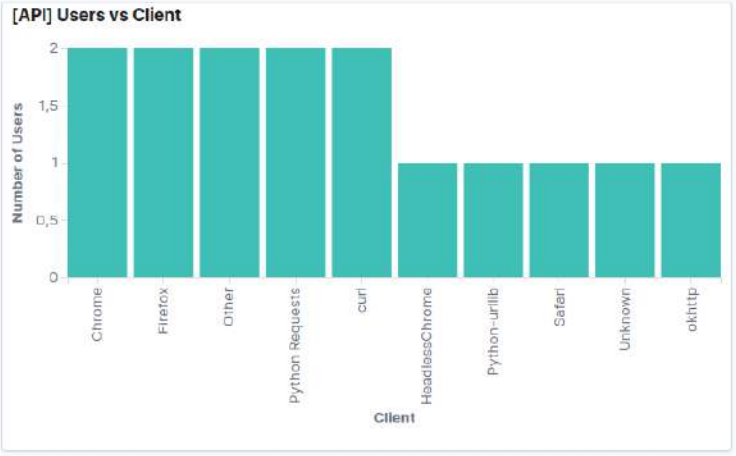




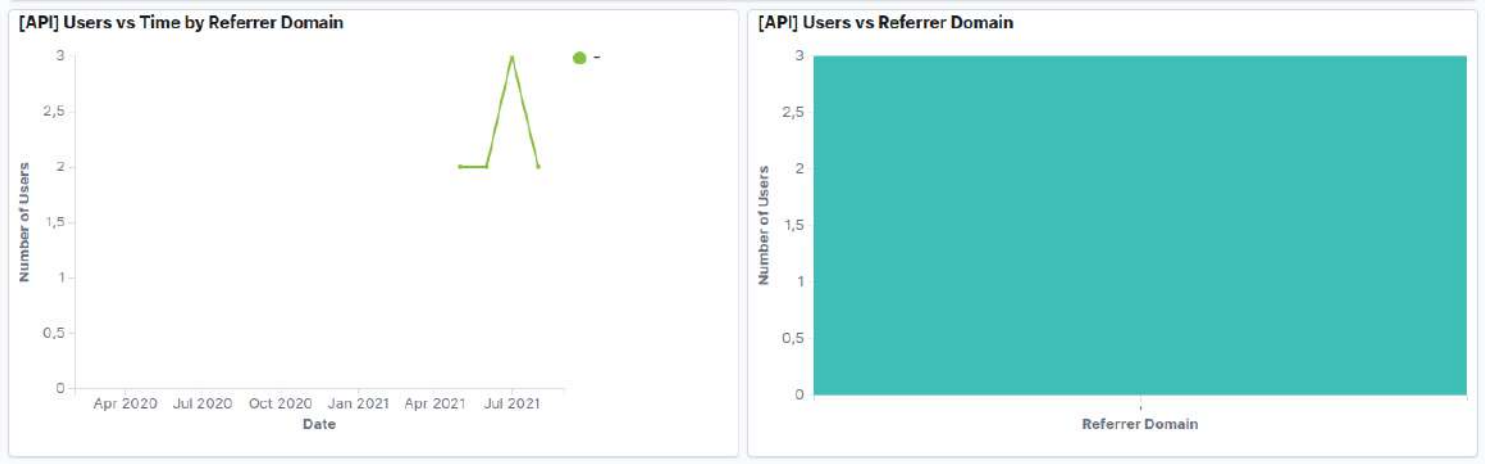




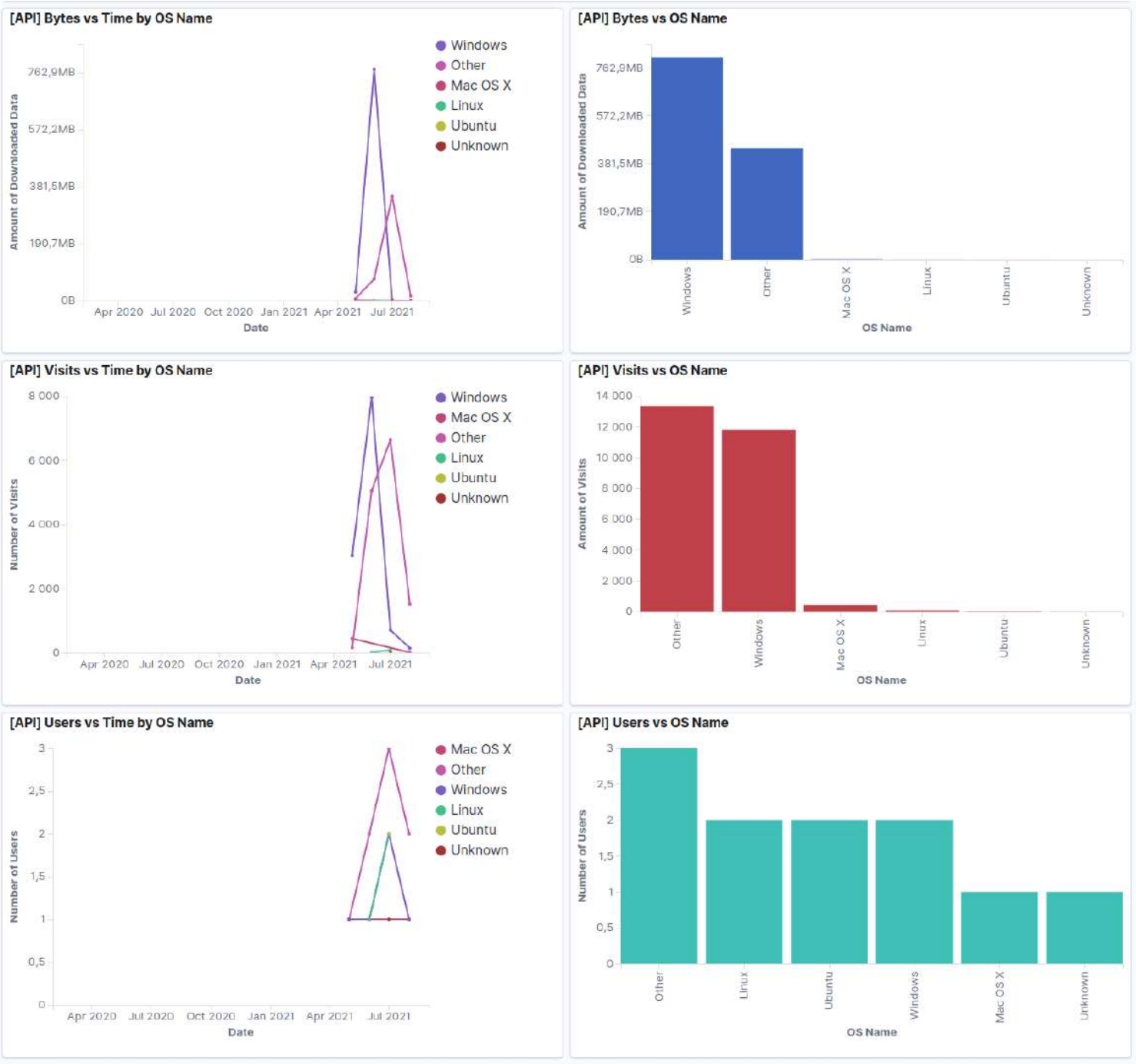




The image below contains an analysis based on the referrer domain ("-" means referrer is not informed, like a visit from a bookmarked link or direct access via software).



Finally, an analysis of traffic based on the operating system of the user is presented (that information is extracted from the browser version, for HTTP libraries, the information is not available and it is contained in the "Other" category).



It can be seen how the "data" endpoints have a bigger proportion of downloaded data than visits. This makes sense, as responses from a "data" endpoint will contain, on average, more bytes than other endpoints. Some features of these metrics that are due to it being under testing can be seen:

- For such a small number of users, there is a wide range of types of software.
- VAMS has not been able to assign a country to most of the IPs the server sees. This indicates that most of the traffic is coming from IPs of the internal SOCIB network

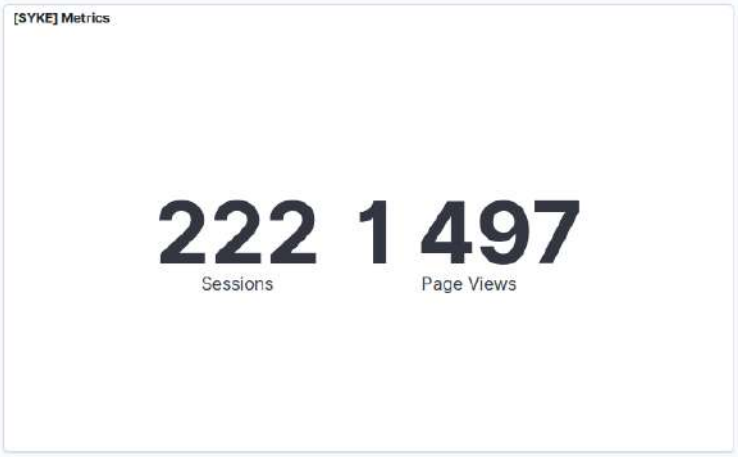
- No user has ever arrived at the API endpoints via a referrer.

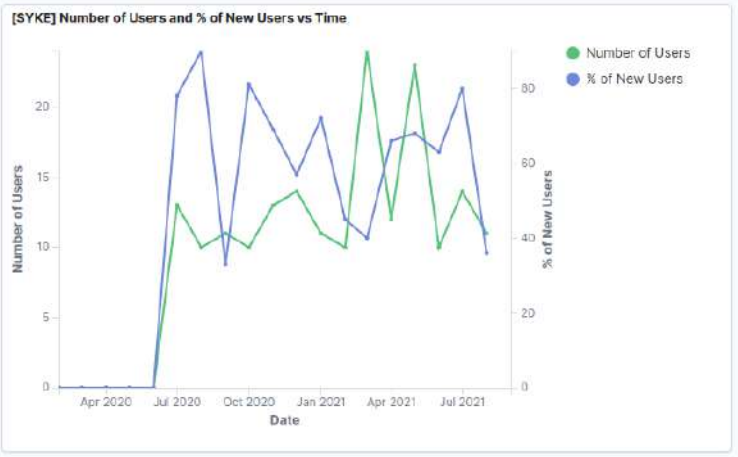
Annexe 1.17 SYKE-ALG@LINE (ID 34.1)

The data underlying the following plots are obtained via a Google Form that SYKE staff kindly fill monthly based on their Google Analytics reports.

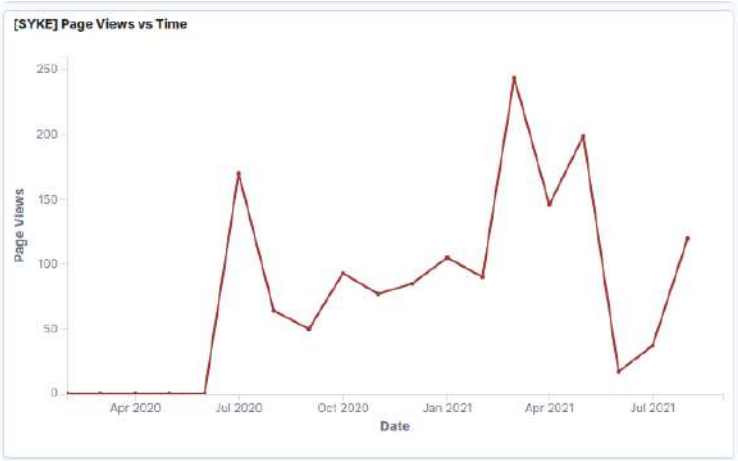
First, the volume of traffic is analysed. Then, metrics related to session quality (session duration and bounce rate) are shown. Finally, the origin of traffic is analysed.

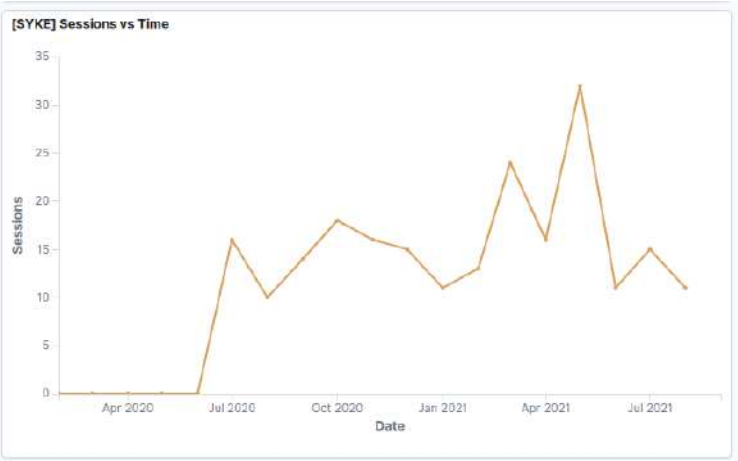
In the next image, a numeric representation of the access metrics and the users analysis are presented.



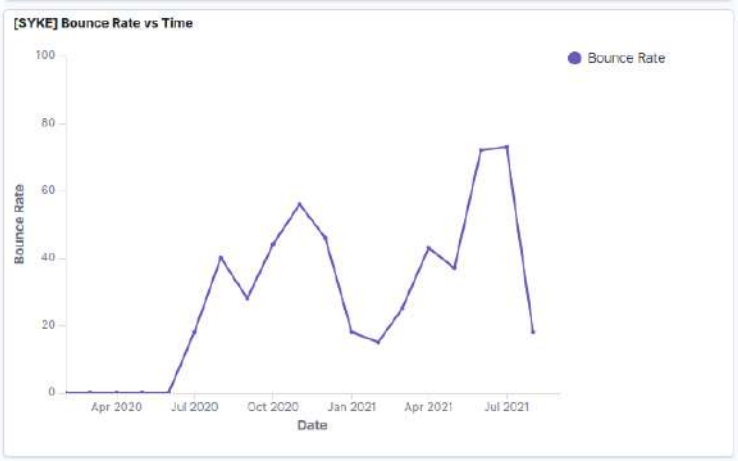


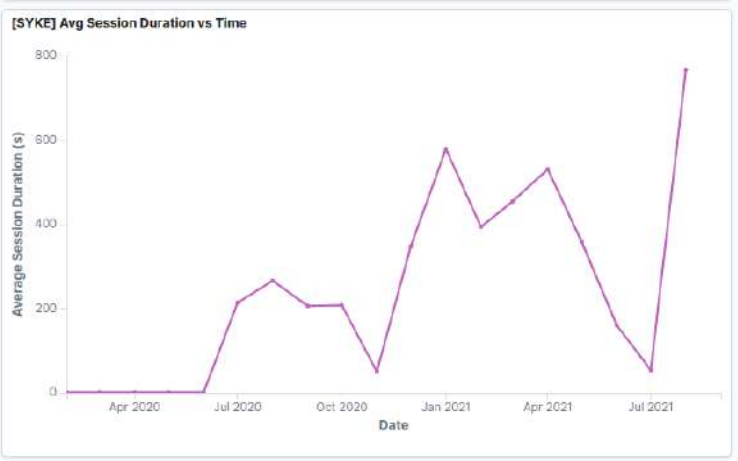
Next, a temporal analysis of the pageviews and sessions have been added.



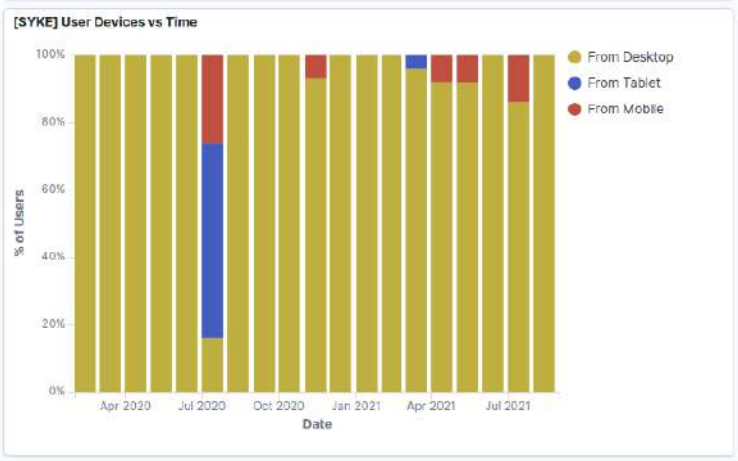


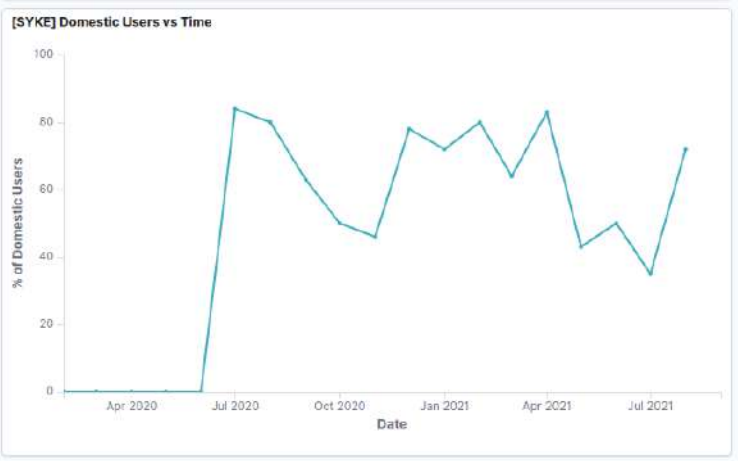
In the following image, metrics related to the on-page user experience: bounce rate and session duration are analysed.





Finally, the user origin is analysed: both by device and country of origin.



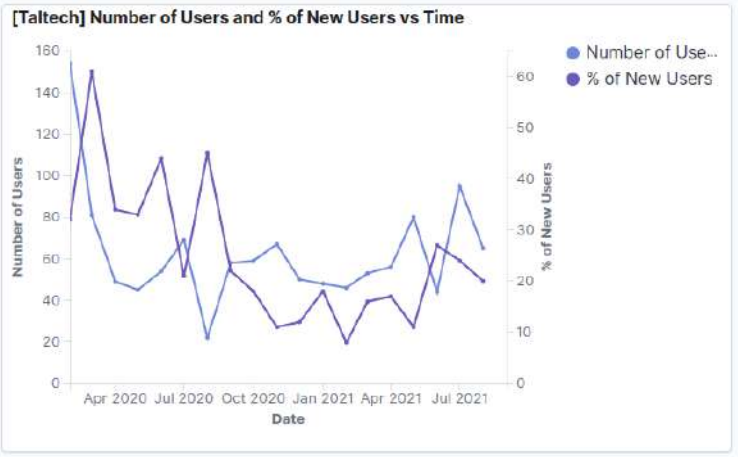


Annexe 1.18 Keri Island Research Station (ID 35.1)

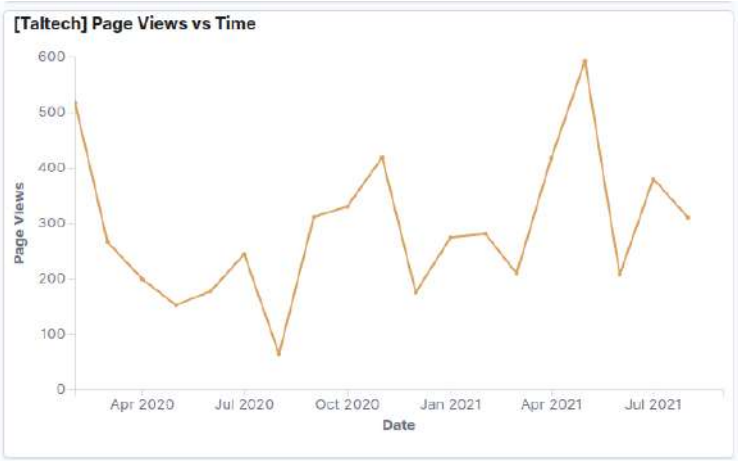
The data underlying the following plots are obtained via a Google Form that Taltech staff kindly fill monthly based on their Google Analytics reports.

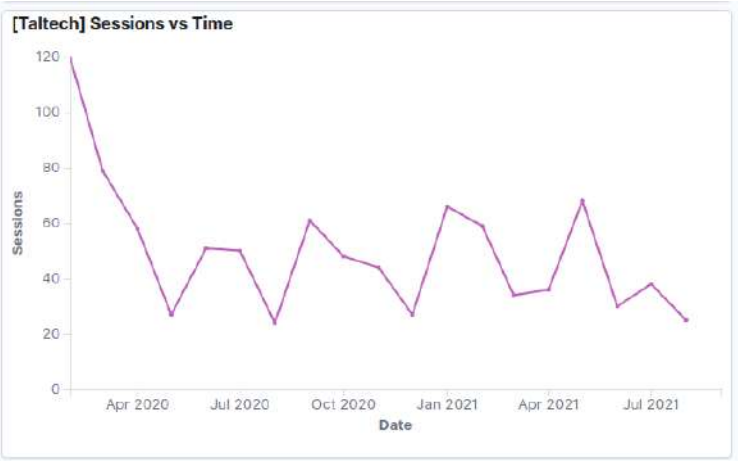
First, the volume of traffic is analysed. Then, metrics related to session quality (session duration and bounce rate) are shown. Finally, the origin of traffic is analysed.

In the next image, a numeric representation of the access metrics and the users analysis are presented.

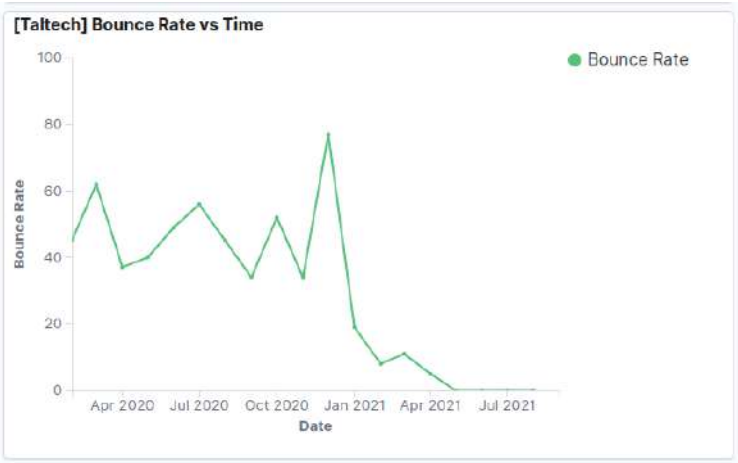


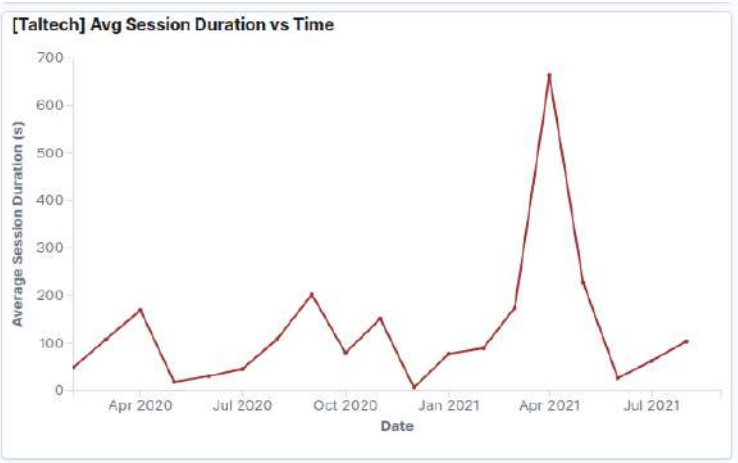
Next, a temporal analysis of the pageviews and sessions have been added.



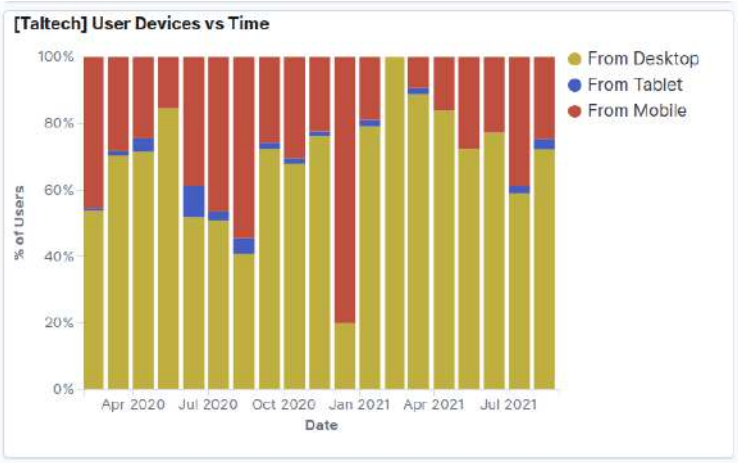


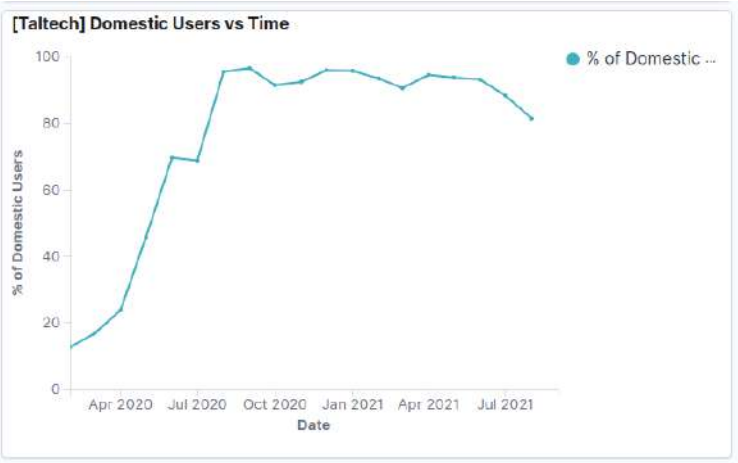
In the following image, metrics related to the on-page user experience: bounce rate and session duration are analysed.





Finally, the user origin is analysed: both by device and country of origin.





Annexe 1.19 OBPS-OTGA (ID 36.1)

This VA infrastructure consists of three parts: AquaDocs, OceanTeacher and OceanBestPractices. All of them are monitored with Google Analytics.

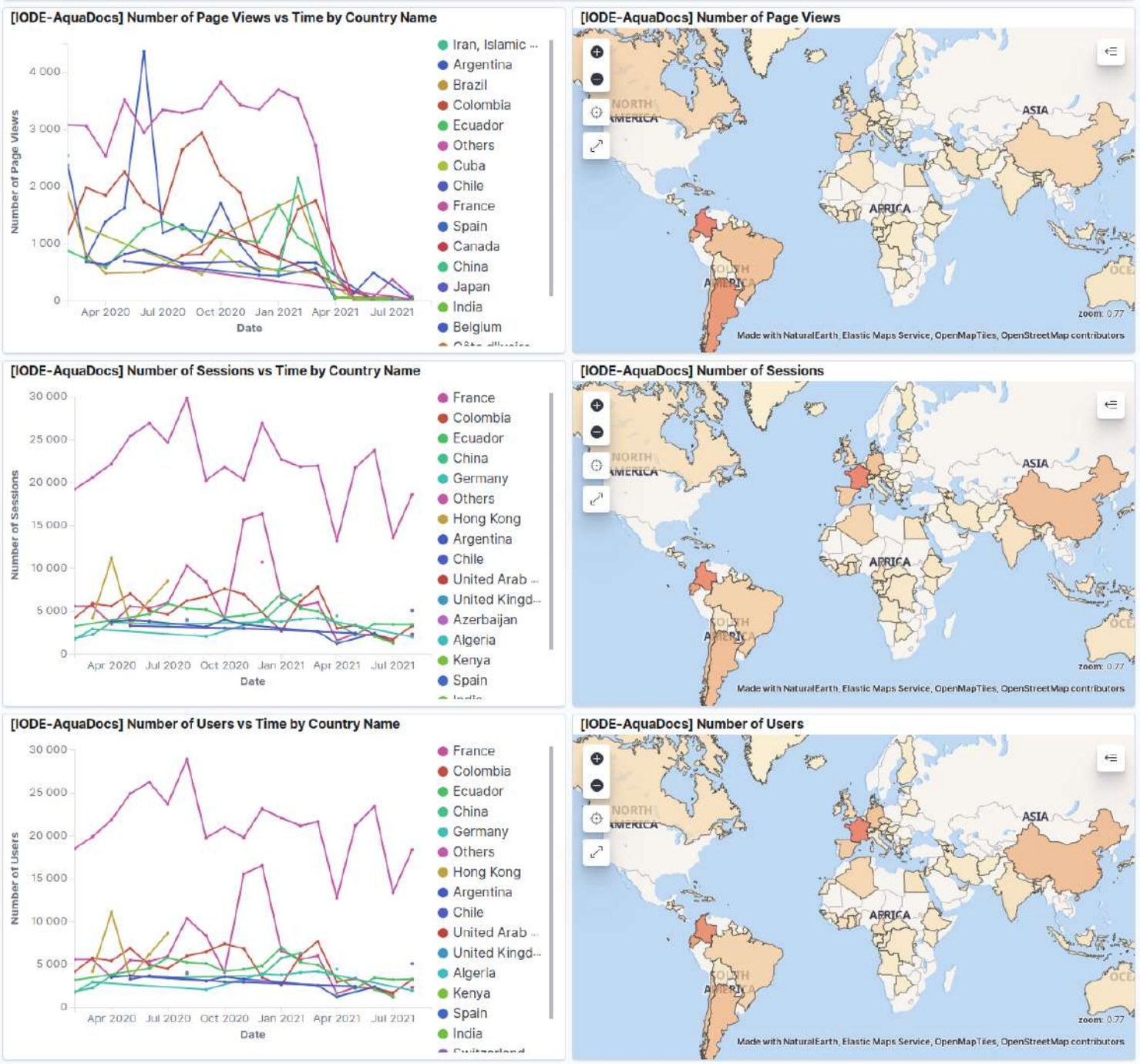
The data underlying all the plots in this annexe have been obtained from the Google Analytics API.

AquaDocs

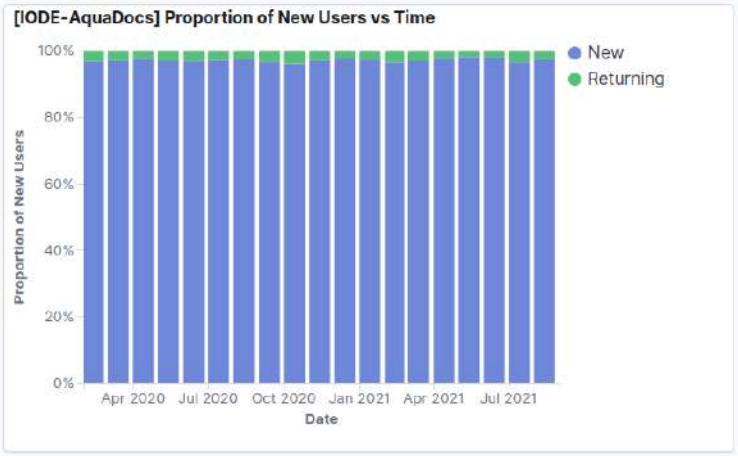
First, the overall access metrics will be presented. Then, the following analysis are included: geographic analysis, analysis by user type (new vs returning), bounce rate analysis, user device analysis and traffic source analysis.

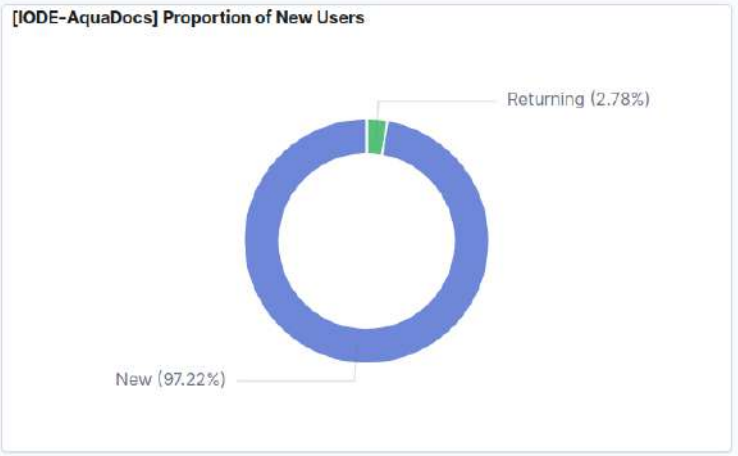
In the next image, the general access metrics in a numeric form are presented.

The following image contains the geographic analysis of the traffic.

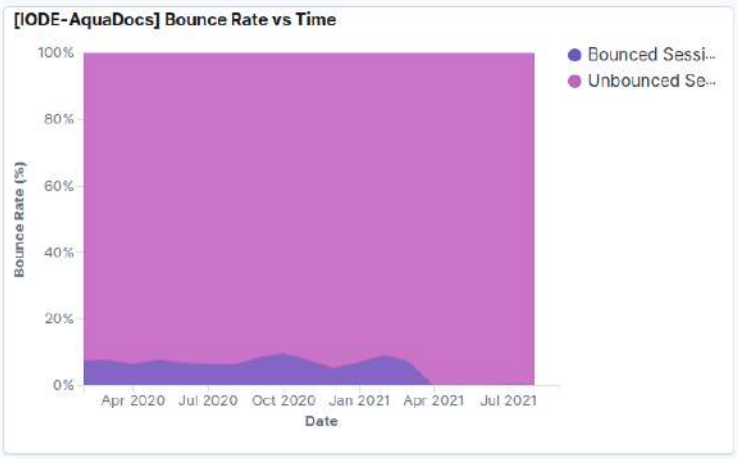


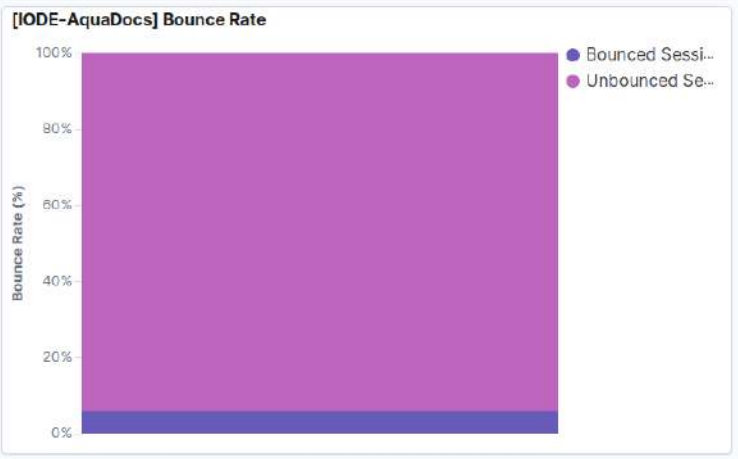
Next, plots analysing the proportion of new users are shown.



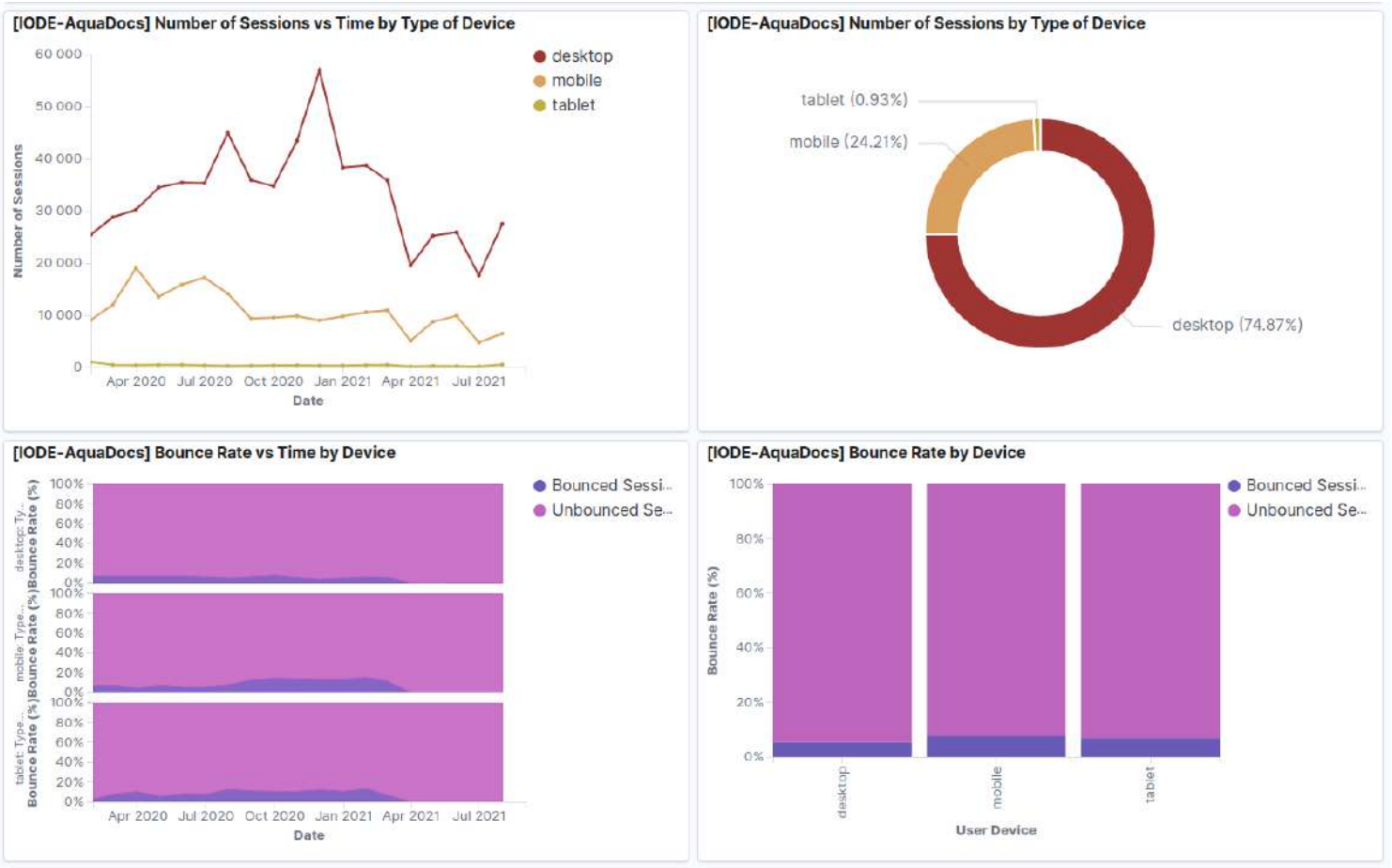


Next, the bounce rate is analysed.

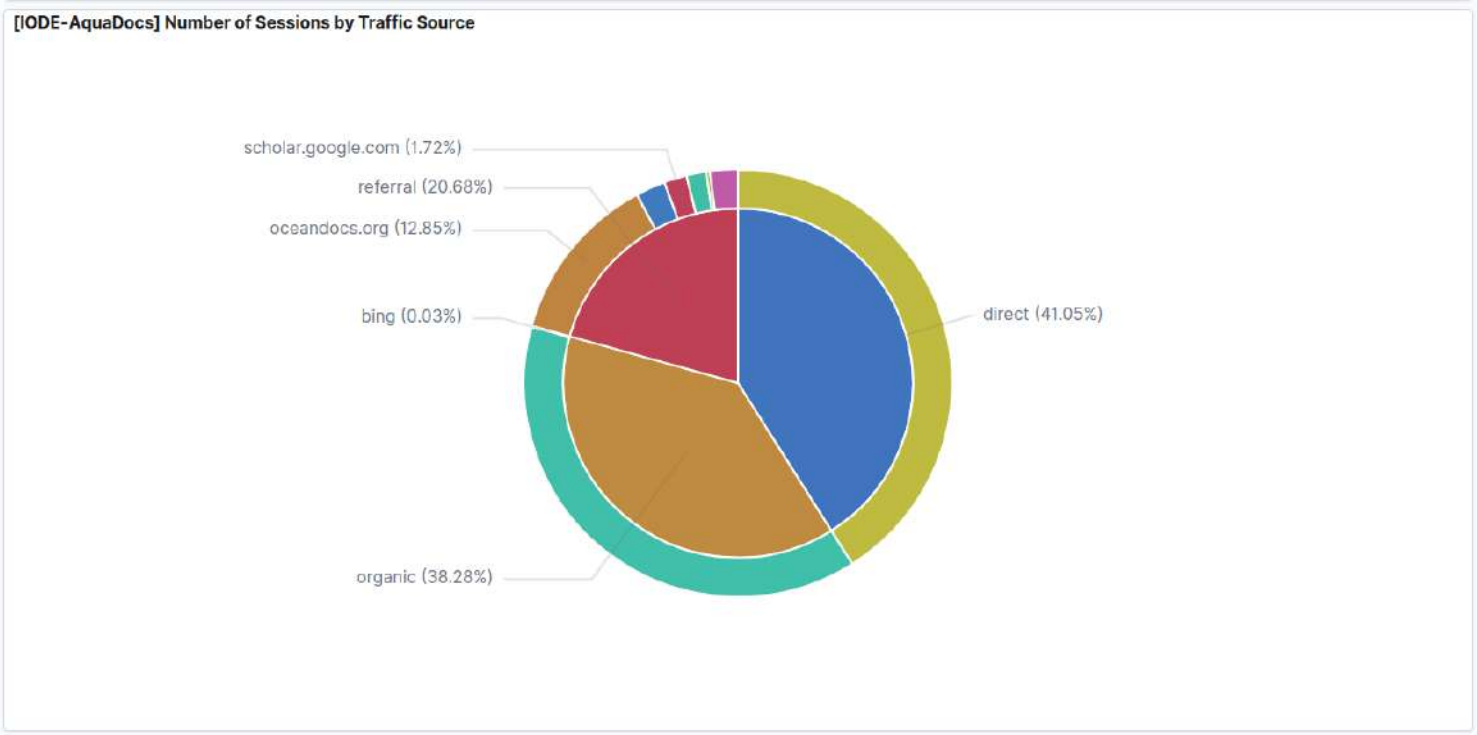




The image below shows the proportion of sessions that come from each type of device and the influence of the type of device on the bounce rate.



Finally, an analysis of traffic based on the source is presented.

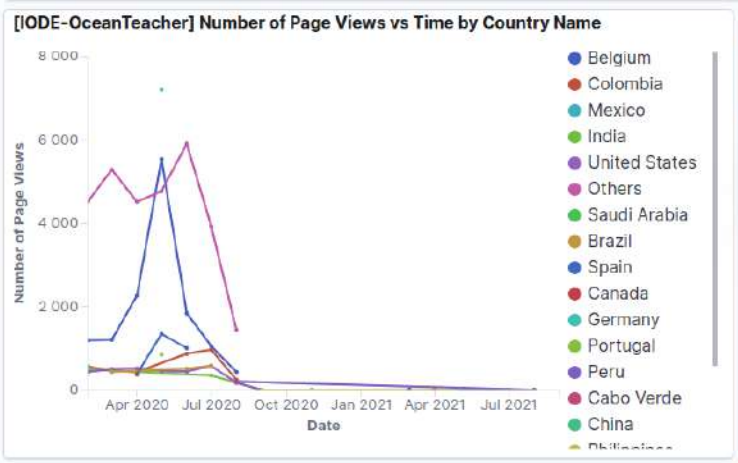


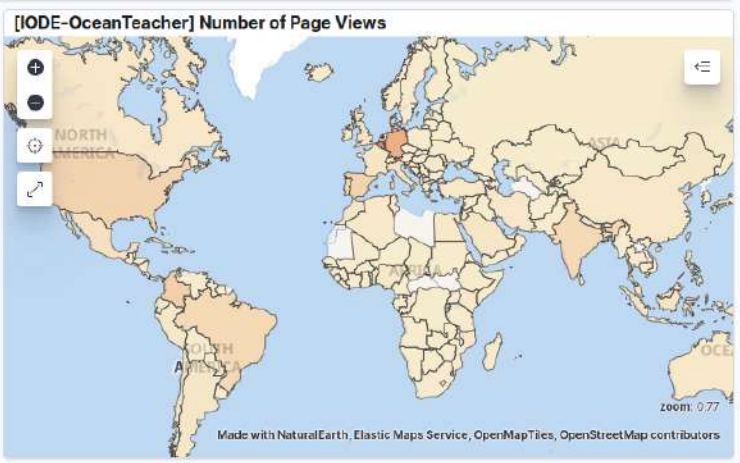
OceanTeacher

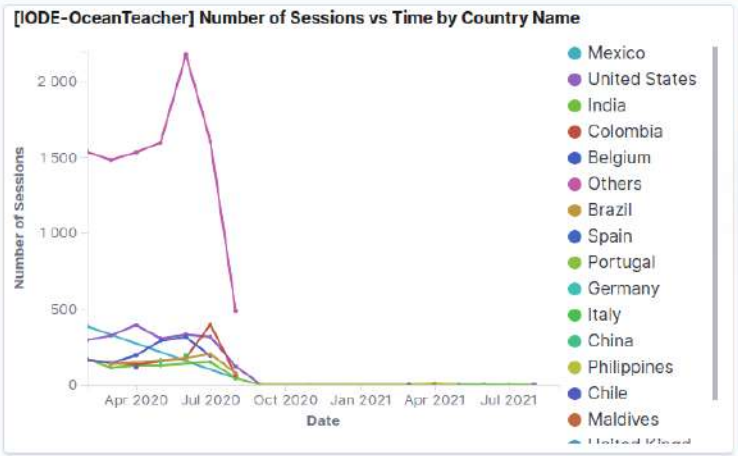
First, the overall access metrics will be presented. Then, the following analysis are included: geographic analysis, analysis by user type (new vs returning), bounce rate analysis, user device analysis and traffic source analysis.

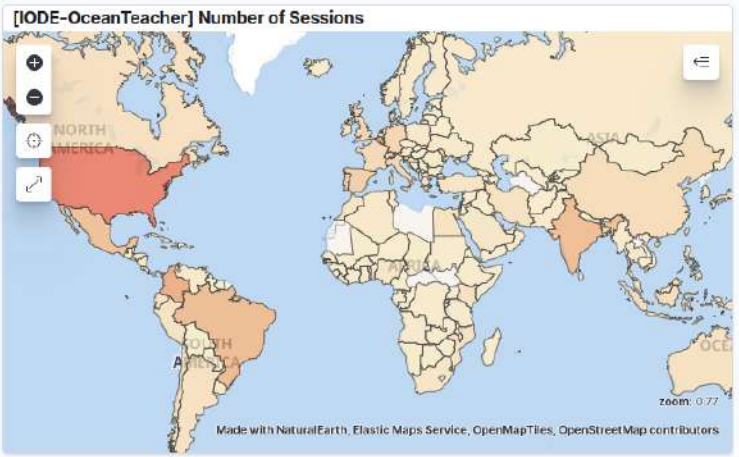
In the next image, the general access metrics in a numeric form are presented.

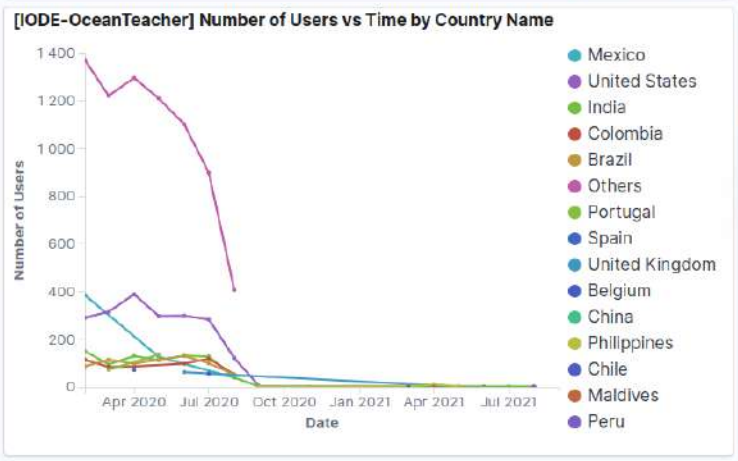
The following image contains the geographic analysis of the traffic.

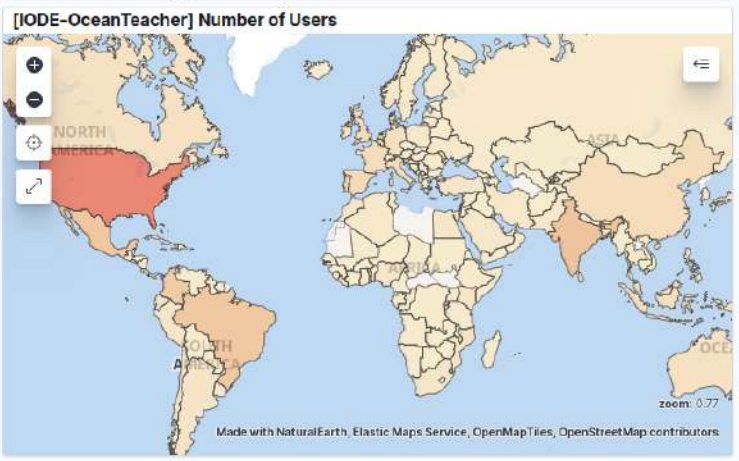




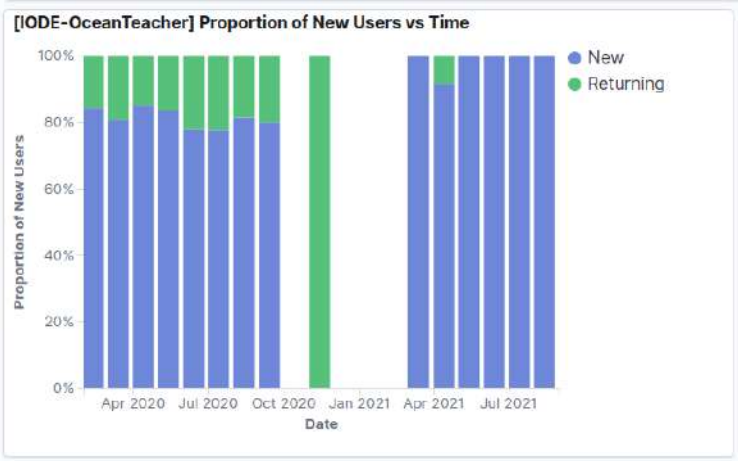


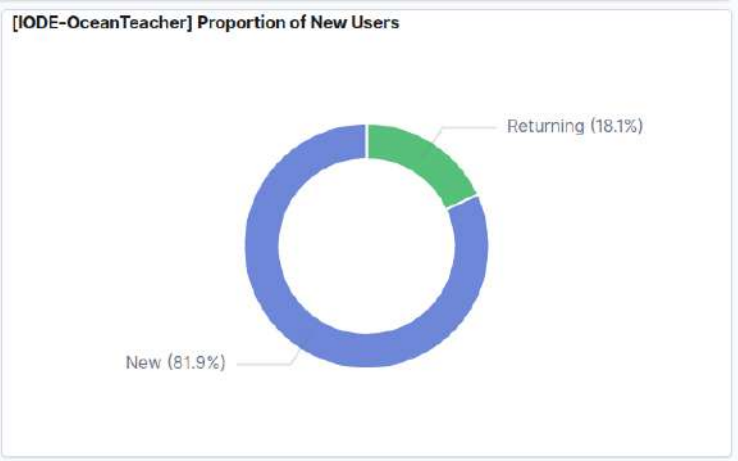




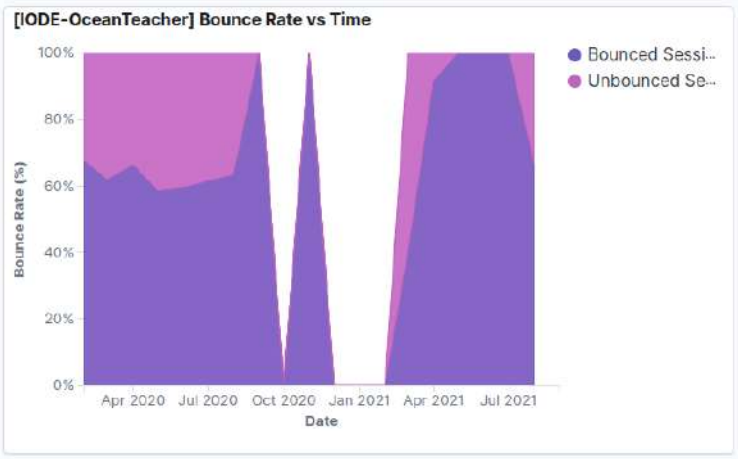


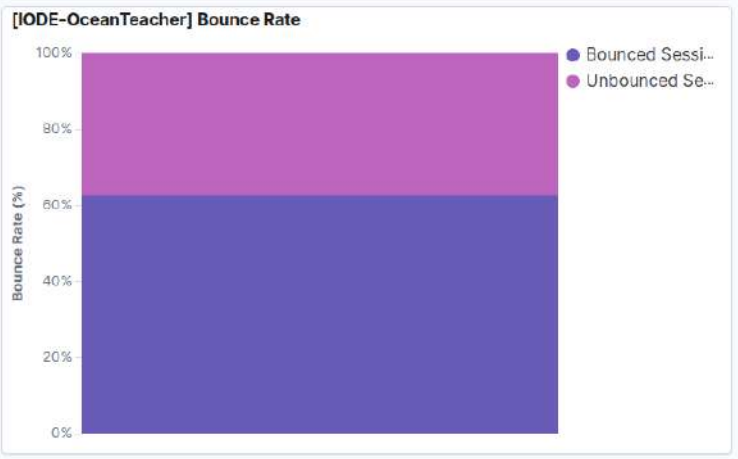
Next, plots analysing the proportion of new users are shown.



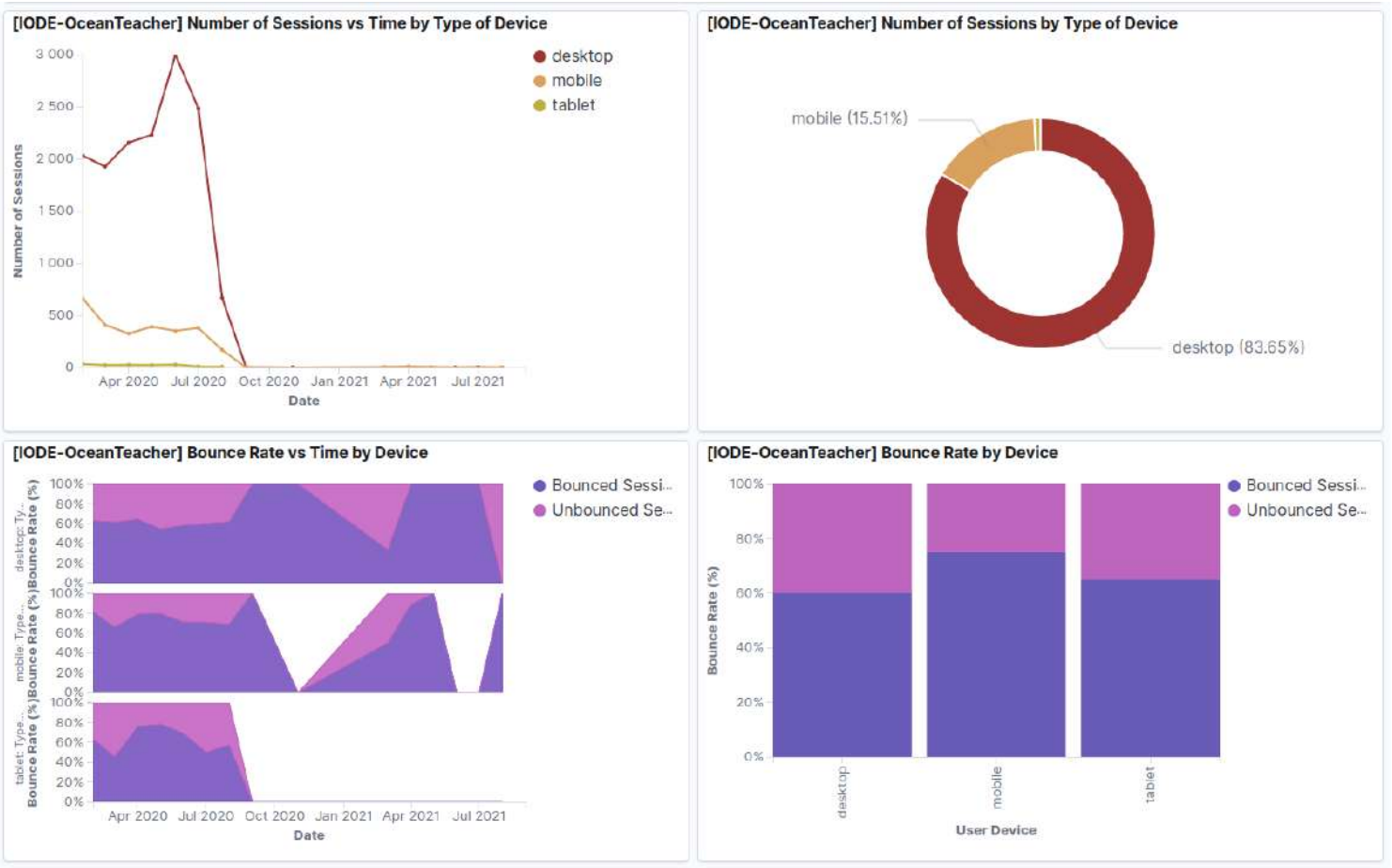


Next, the bounce rate is analysed.

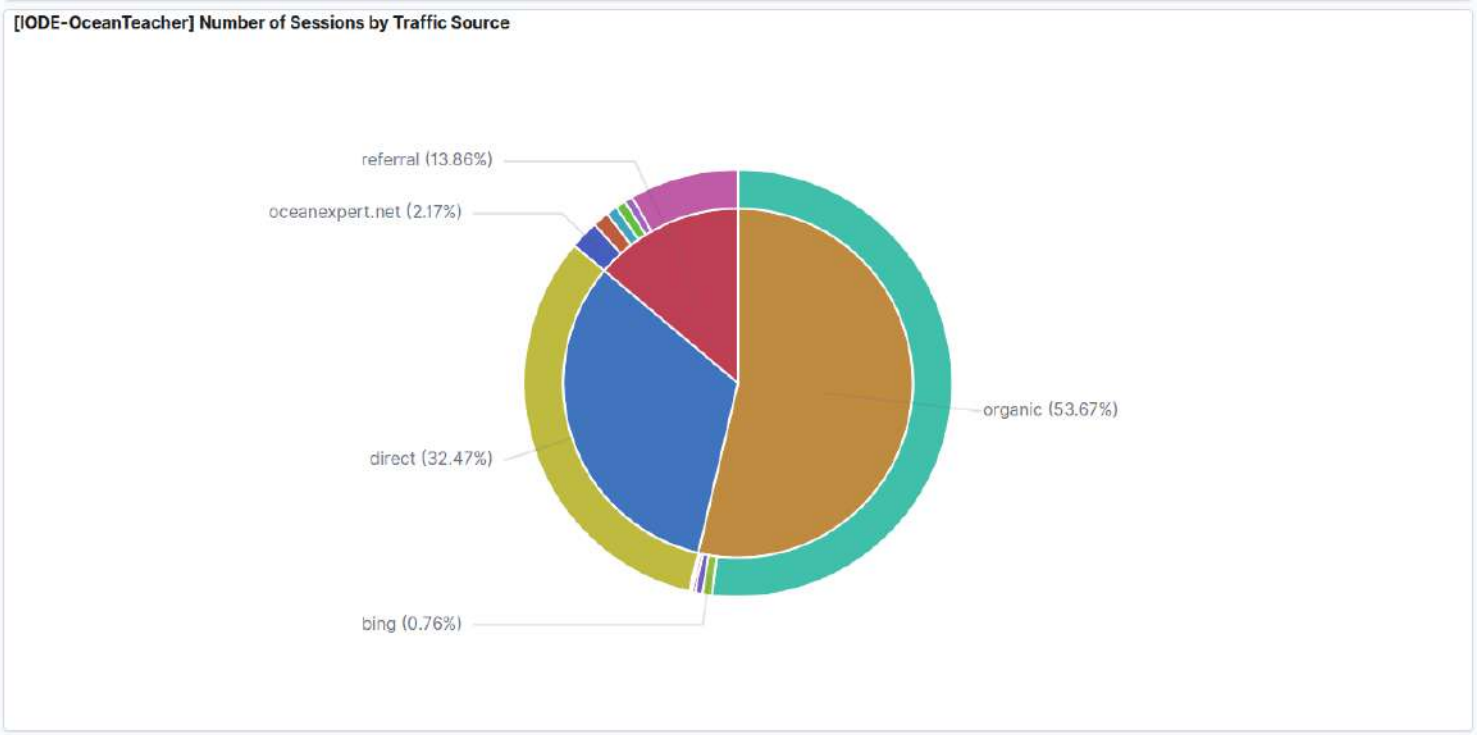




The image below shows the proportion of sessions that come from each type of device and the influence of the type of device on the bounce rate.



Finally, an analysis of traffic based on the source is presented.

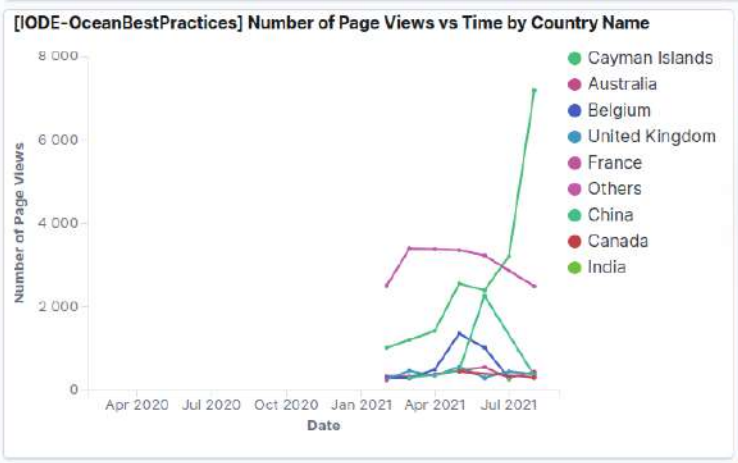


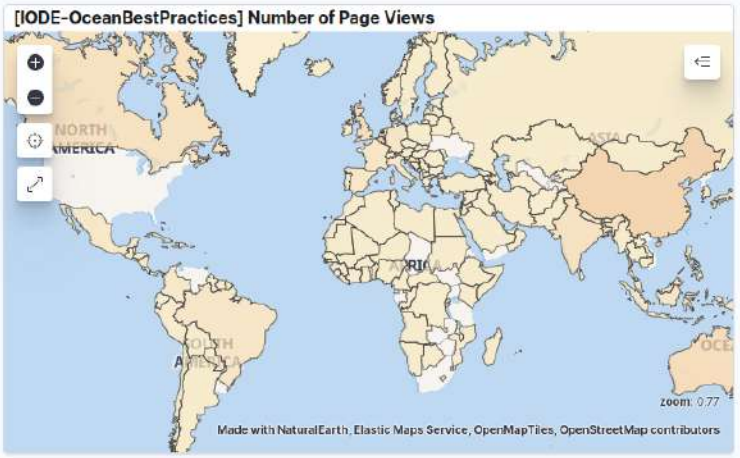
OceanBestPractices

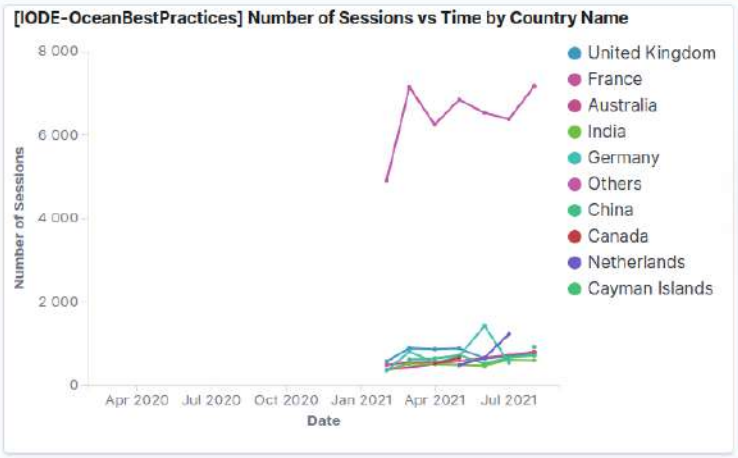
First, the overall access metrics will be presented. Then, the following analysis are included: geographic analysis, analysis by user type (new vs returning), bounce rate analysis, user device analysis and traffic source analysis.

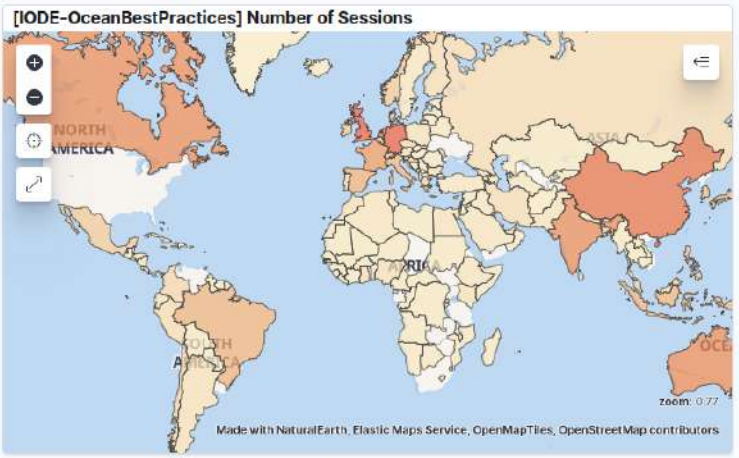
In the next image, the general access metrics in a numeric form are presented.

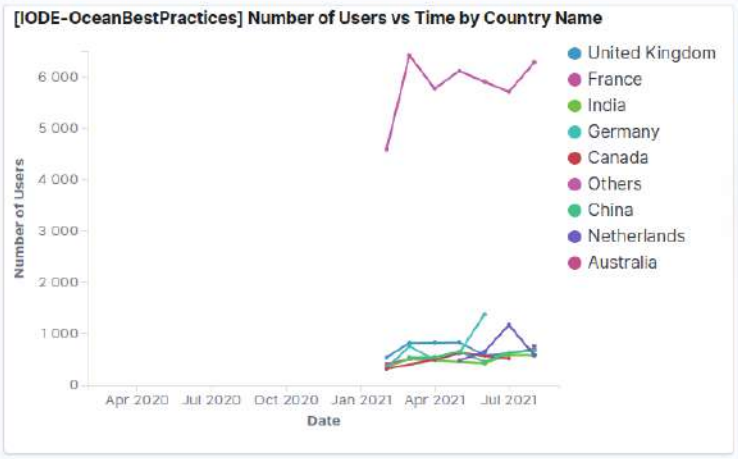
The following image contains the geographic analysis of the traffic.

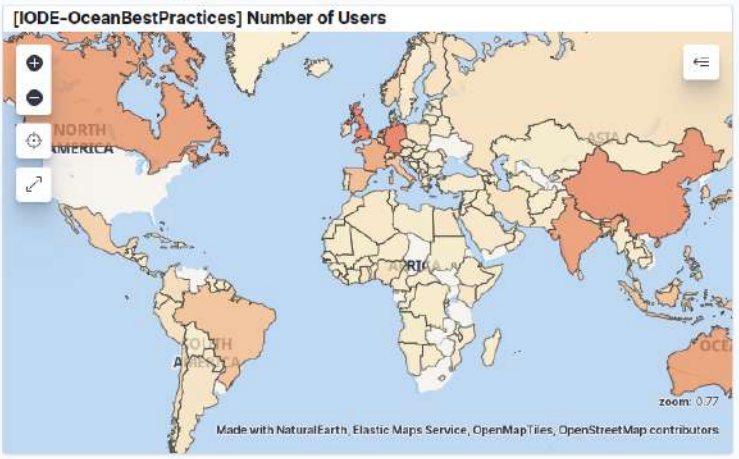




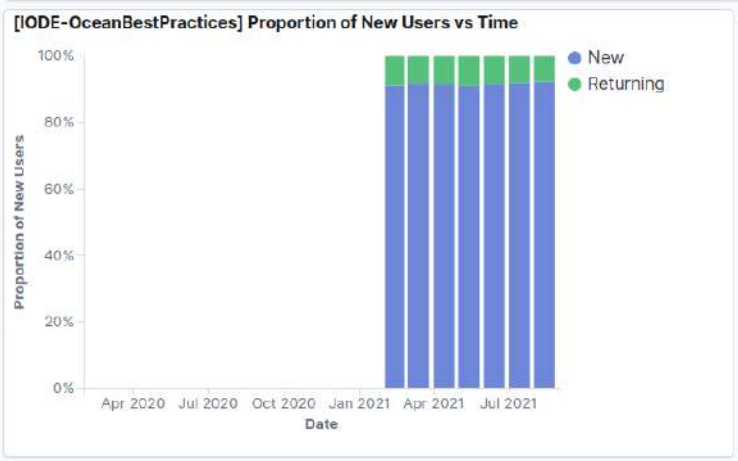


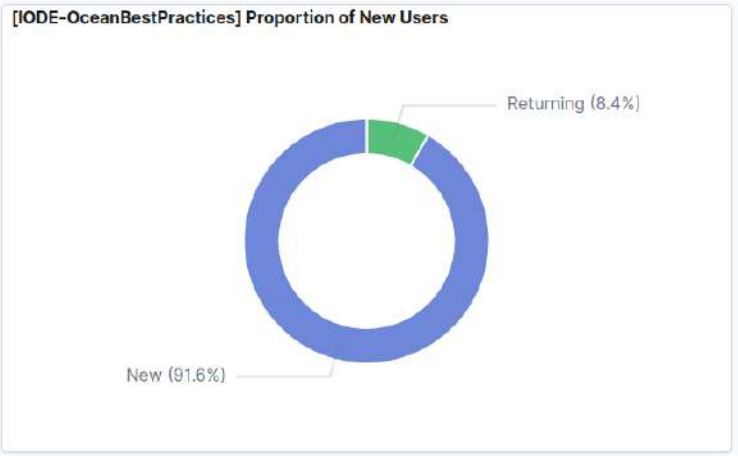




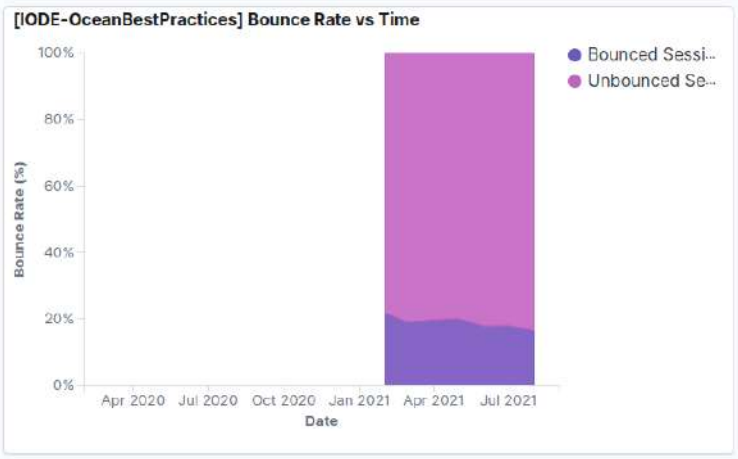


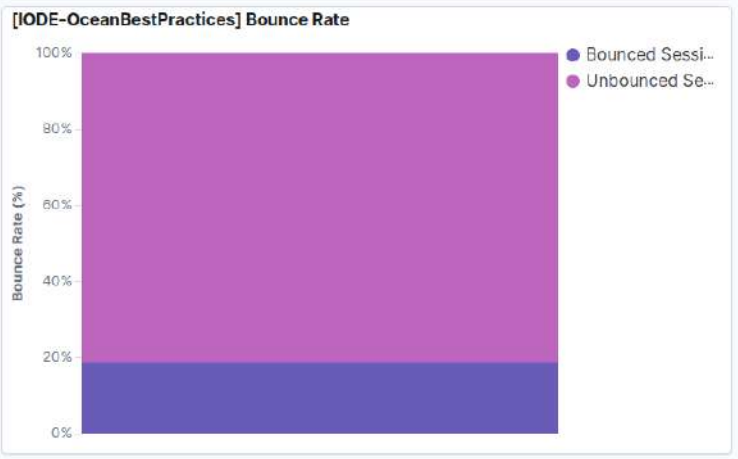
Next, plots analysing the proportion of new users are shown.





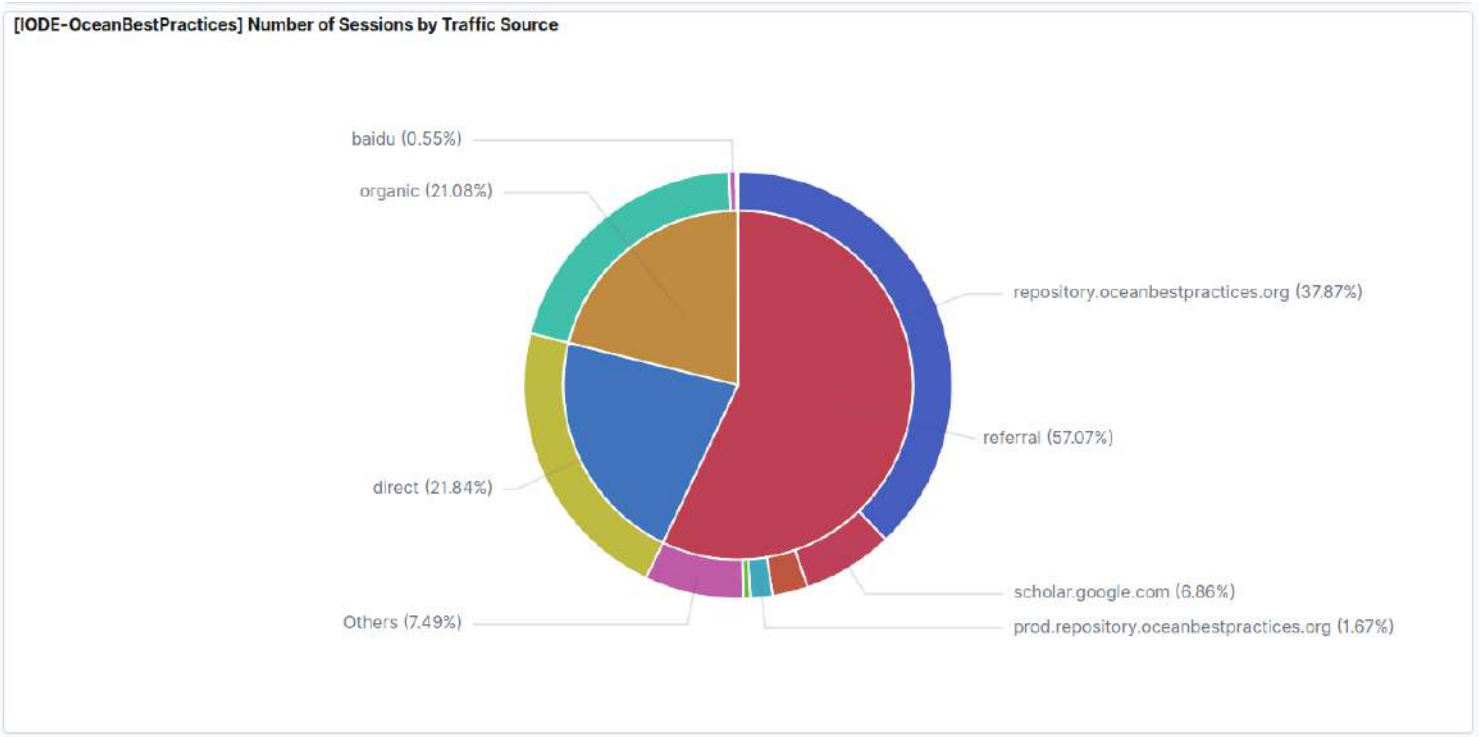
Next, the bounce rate is analysed.





The image below shows the proportion of sessions that come from each type of device and the influence of the type of device on the bounce rate.

Finally, an analysis of traffic based on the source is presented.

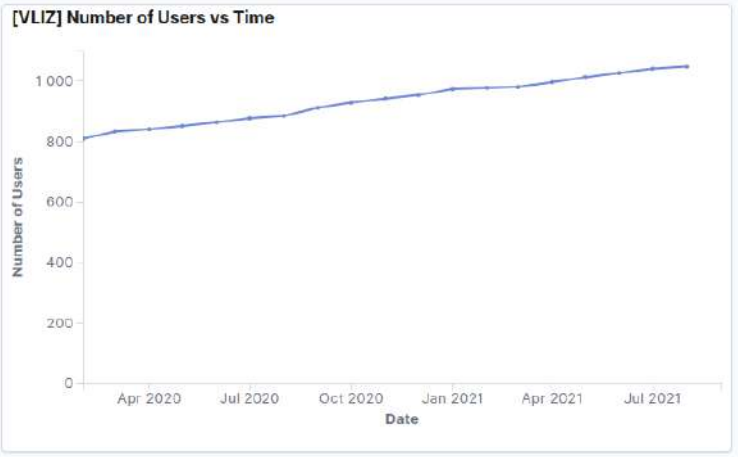


Annexe 1.20 VLIZ Marine Data Archive (ID 38.1)

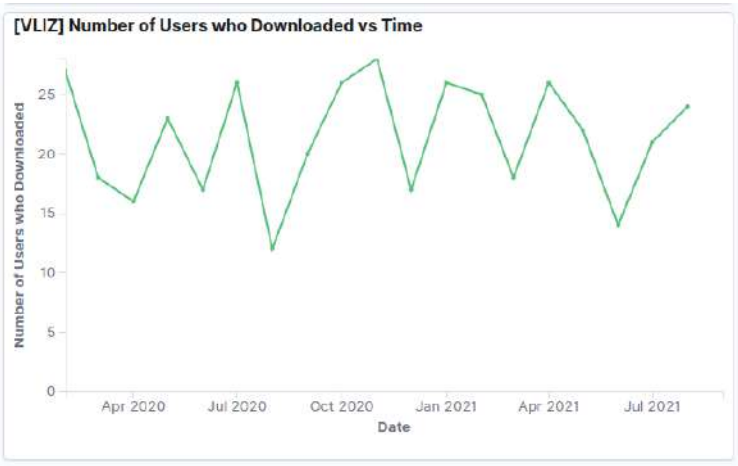
The data underlying the following plots are kindly generated by VLIZ's team and made available in a CSV file.

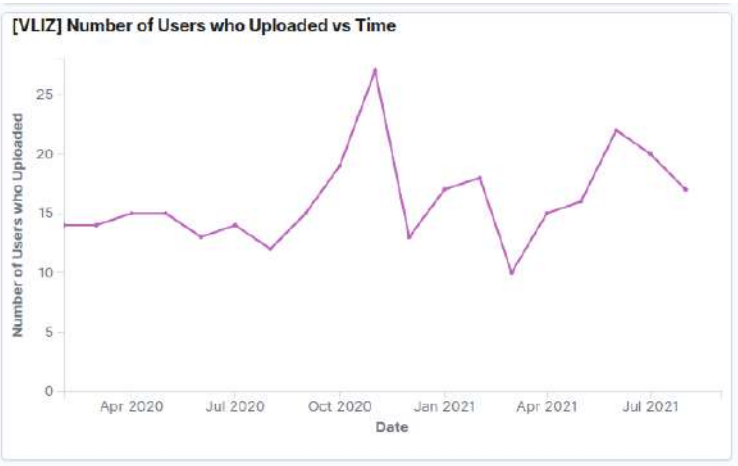
First, overall metrics and a temporal trend of the number of users will be presented. Next, the user activity analysis, the uploads analysis and the download analysis will be analysed.

In the next image, a numeric representation of the access metrics and the temporary trend of the number of users is presented.

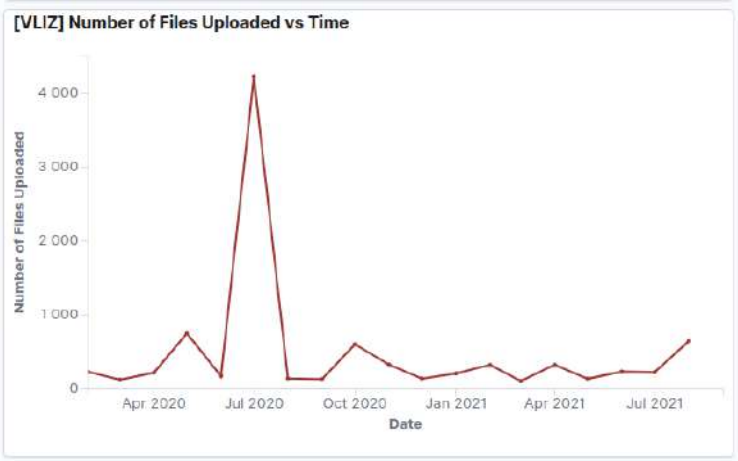


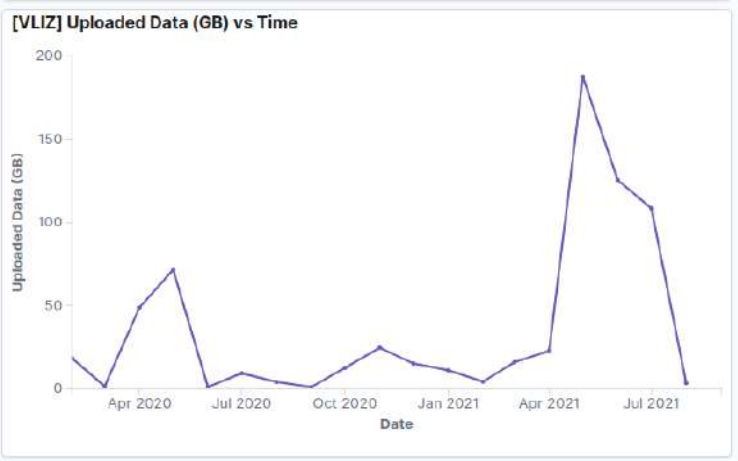
In the following image, a temporal analysis of the number of users who upload or download data has been added.



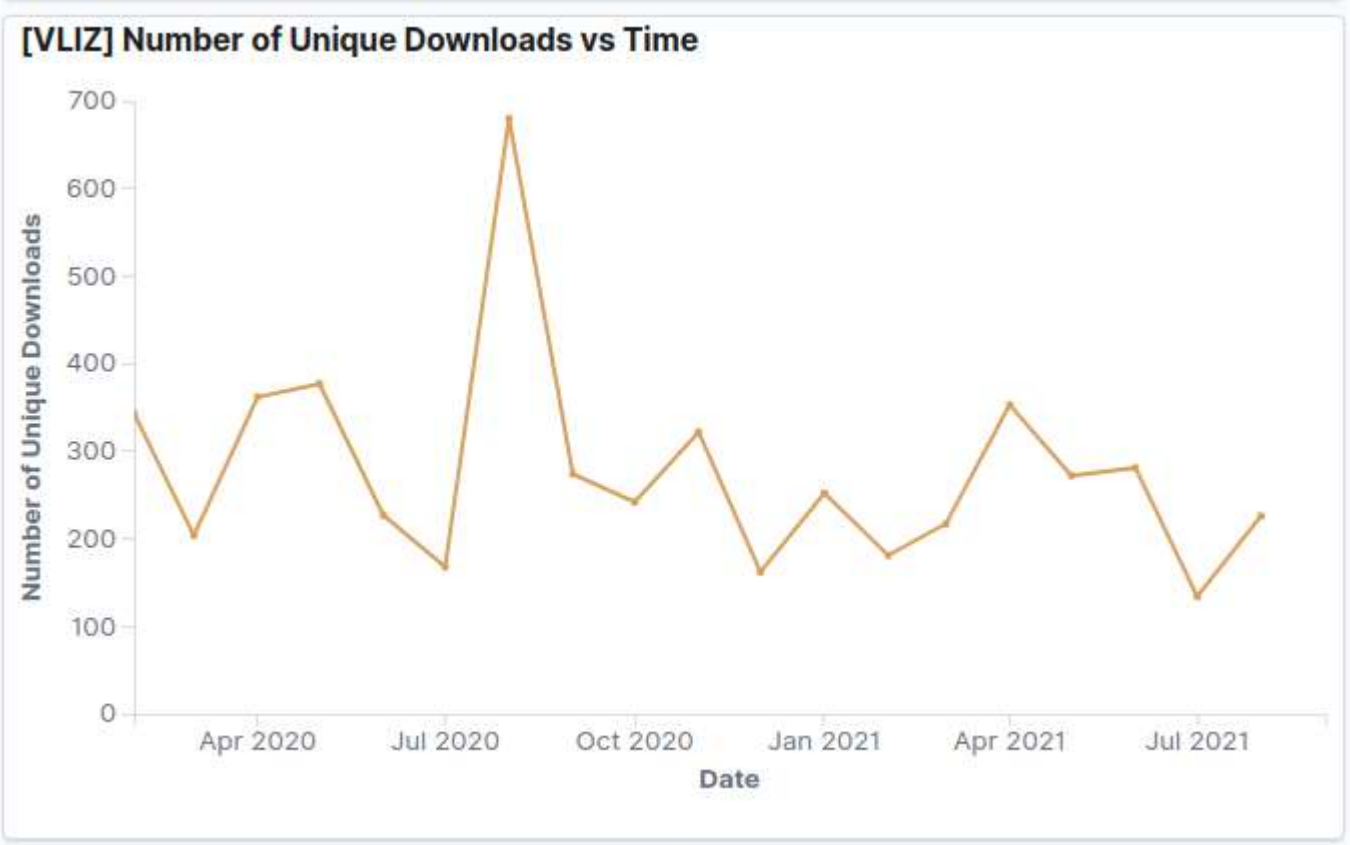


Next, the uploads both by number and by size are analysed.





Finally, downloads by number are analysed.



Annexe 2: Outreach Activities from each Partner

The outreach activities of JERICO-S3 WP11 are reported next.

Annexe 2.1 Mawenzi (ID 1.2)

The following is the list of outreach activities reported by Mawenzi, which belongs to IFREMER ULCO-LISIC.

Title	uHMM documentation in French and English
Date	4 May 2021
Activity	Type: Website. Channel: Website
Audience	Type: Scientific community, Industry. Size: large
URL	http://mawenzi.univ-littoral.fr/uHMMweb/

Title	sClust tools : documentation, shiny app and R package - multi-level spectral Clustering algorithms
Date	19 July 2021
Activity	Type: Website. Channel: Website, gitlab
Audience	Type: Scientific community, Industry. Size: Large
URL	http://mawenzi.univ-littoral.fr/sClust/

Title	ULC0-LISIC poster seminar : presentation of 3 posters about LISIC contribution for JERICO project
Date	6 July 2021
Activity	Type: Participation in an event other than a conference or workshop. Channel: ULCO Lab audience
Audience	Type: Scientific community, Students. Size: 70

Annexe 2.2 EU HFR Node/AZTI (ID 4.1)

The following is the list of outreach activities reported by EU HFR Node/AZTI, which belongs to AZTI.

Title	Emails to request data and offer EU HFR Node processing to data provider
Date	17 May 2021
Activity	Type: Data request from EU HFR Node. Channel: Email and newsletter
Audience	Type: Scientific community. Size: 1 data provider per email (4 emails sent)

Title	Oral presentation of the EU HRF Node Historical Data processing
Date	27 May 2021
Activity	Type: Training. Channel: Online HF Radar Summer School
Audience	Type: Scientific community. Size: 50
URL	https://sicomarschool.univ-tln.fr/en/homepage/

Title	EuroGOOS HFR Task Team's 3rd biannual newsletter.
Date	19 July 2021
Activity	Type: Communication Campaign (e.g. Radio or TV). Channel: Website, Email and newsletter
Audience	Type: Scientific community, Industry, Investors, Customers. Size: 100
URL	https://us19.campaign-archive.com/?e=__test_email__&u=e5fd08d8d94228eecb45183f9&id=1de066b71f

Annexe 2.3 CefMAT (ID 6.1)

The following is the list of outreach activities reported by CefMAT, which belongs to Cefas.

Date	Year 2020
Activity	Type: Website. Channel: Website
URL	https://www.ulsterwildlife.org/sustainable-fish-cities-ni

Date	Year 2020
Activity	Type: Website. Channel: Website
URL	https://data.catchmentbasedapproach.org/apps/theriverstrust::policy-legislation/explore

Date	Year 2020
Activity	Type: Website. Channel: Website
URL	https://www.daera-ni.gov.uk/articles/marine-strategy

Title	Department for Environment, Food and Rural Affairs.
Date	24 September 2020
Activity	Type: Organisation of a workshop. Channel: Public Debate

Title	Plastics: Pollution. Department for Environment, Food and Rural Affairs.
Date	12 November 2020
Activity	Type: Organisation of a workshop. Channel: Public Debate

Annexe 2.4 EU HFR Node/CNR (ID 7.2)

The following is the list of outreach activities reported by EU HFR Node/CNR, which belongs to CNR.

Title	Sicomar-Plus HF radars summer school
Date	27 May 2021
Activity	Type: Participation in an event other than a conference or workshop. Channel: Summer School
Audience	Type: Scientific community, Industry, Civil society, Customers. Size: 50
URL	https://sicomarschool.univ-tln.fr/en/summer-school-sicomar/

Annexe 2.5 CytoFluoTool (ID 8.1)

The following is the list of outreach activities reported by CytoFluoTool, which belongs to CNRS-LOG, ULCO.

Title	ULC0-LISIC poster seminar : presentation of 3 posters about LISIC contribution for JERICO project
Date	6 July 2021
Activity	Type: Participation in an event other than a conference or workshop. Channel: ULCO Lab audience
Audience	Type: Scientific community, Students. Size: 70

Title	Automated approaches for phytoplankton monitoring, at high resolution, in coastal waters: advantages and challenges
Date	23 November 2020
Activity	Type: Participation in a workshop. Channel: Zoom
Audience	Type: Scientific community, Policymakers, Harbour and coastal engineering community. Size: 50 persons

Title	Phytoplankton in vivo/in situ observations by novel automated optical approaches in coastal and marine systems: towards a better integration into joint observatories
Date	5 May 2021
Activity	Type: Participation in a workshop. Channel: Zoom platform
Audience	Type: Scientific community, Policymakers, Harbour and coastal engineering community. Size: 100

Title	- ASLO International Aquatic Sciences Meeting (ASM) 2021: Special Session 66 on Coastal Ocean Observing Systems to understand and predict changes of the coastal ocean
Date	27 June 2021
Activity	Type: Organisation of a conference. Channel: Zoom
Audience	Type: Scientific community. Size: 50

Title	ASLO International Aquatic Sciences Meeting (ASM) 2021: Special Session 28 on Aquatic microbial community structure and dynamics: new insights from non-destructive high throughput automated single-cell analysis
Date	23 June 2021
Activity	Type: Organisation of a conference. Channel: Zoom platform
Audience	Type: Scientific community. Size: 75

Title	ASLO International Aquatic Sciences Meeting (ASM) 2021: Special Session 28 on Aquatic microbial community structure and dynamics: new insights from non-destructive high throughput automated single-cell analysis
Date	19 March 2021
Activity	Type: Training. Channel: In person
Audience	Type: Students. Size: 20

Title	M.Sc. on Marine Sciences: Marine Ecology and Fisheries at University of Littoral: "Automated tools and approaches for marine observations"
Date	25 March 2021
Activity	Type: Training. Channel: In person
Audience	Type: Students. Size: 20

Annexe 2.6 JERICO-ECOTAXA (ID 8.2)

The following is the list of outreach activities reported by JERICO-ECOTAXA, which belongs to CNRS-LOV.

Title	Plankton and marine snow dynamics inferred from their morphological attributes : global context of the BCP, case study in the Arctic, perspectives Emilia Trudnowska , Lars Stemmann Seminar of JETZON - Joint Exploration of the Twilight Zone Ocean Network, 28/05/2021
Date	28 May 2021
Activity	Type: Participation in a conference. Channel: internet
Audience	Type: Scientific community. Size: 40

Title	Typologie globale des spectres de taille du zooplancton dans le premier kilometre des océans Looking for global relationships. Dodji Soviadan, Baye Cheickh Mbaye, Laetitia Drago, Lars Stemmann, Seminar of Third meeting of the ANR project CIGOEF 2 and 10 June 2021
Date	10 February 2021
Activity	Type: Participation in a conference. Channel: on internet visioconference
Audience	Type: Scientific community. Size: 30

Title	I/ITAPINA: Imagine/Imaging The Atlantic – A Pelagic Imaging Network Approach (I/ITAPINA) workshop to be held on the 28th and 29th June 2021. a AA-MARINET pilot action and supported by the All-Atlantic Ocean Research Alliance
Date	28 June 2021
Activity	Type: Participation in activities organised jointly with other H2020 projects. Channel: Website, Email and newsletter
Audience	Type: Scientific community. Size: 140

Annexe 2.7 Utö Atmospheric and Marine Research Station (ID 15.1)

The following is the list of outreach activities reported by Utö Atmospheric and Marine Research Station, which belongs to FMI.

Title	New data accumulation service
Date	2 April 2020
Activity	Type: Press release. Channel: Website, Email and newsletter
Audience	Type: Scientific community, Industry, Civil society, General public, Policymakers, Media, Investors. Size: >100 000
URL	https://www.ilmatieteenlaitos.fi/ajankohtaista/1269839400

Title	Press release
Date	4 June 2020
Activity	Type: Press release. Channel: Website, Email and newsletter
Audience	Type: Scientific community, Industry, Civil society, General public, Policymakers, Media. Size: > 100 000
URL	https://en.ilmatieteenlaitos.fi/news/4rbaC4TjrITEUFEHVngUzO

Title	Launch of "marinefinland.fi"
Date	4 June 2020
Activity	Type: Website. Channel: Website
Audience	Type: Scientific community, Industry, Civil society, General public, Policymakers, Media, Investors, Customers. Size: > 1 000 000
URL	https://www.marinefinland.fi/en-US

Title	HAB situation
Date	12 August 2021
Activity	Type: Website. Channel: Website

Audience	Type: Scientific community, Industry, Civil society, General public, Policymakers, Media. Size: not applicable

Annexe 2.8 POSEIDON Multi Platform Observatory Data Center (ID 16.1)

The following is the list of outreach activities reported by POSEIDON Multi Platform Observatory Data Center, which belongs to HCMR.

Title	News Announcement for the participation in Jerico S3
Date	21 February 2020
Activity	Type: Website. Channel: Website
Audience	Type: Scientific community, General public. Size: 500
URL	https://poseidon.hcmr.gr/news/jerico-s3-kick-meeting-san-sebastian-spain-february-2020

Title	Launch of new poseidon.hcmr.gr website
Date	1 February 2021
Activity	Type: Website. Channel: Website
Audience	Type: Scientific community, General public. Size: > 1.000.000
URL	https://poseidon.hcmr.gr

Title	Presentation at 18th E-SURFMAR Expert Team Meeting
Date	27 May 2021
Activity	Type: Participation in a conference. Channel: Project Meeting
Audience	Type: Scientific community. Size: 100

Title	Marine Heat Wave in the Aegean Sea in June 2021
Date	6 July 2021
Activity	Type: Press release. Channel: Website
Audience	Type: Scientific community, General public. Size: 200
URL	https://poseidon.hcmr.gr/news/marine-heat-wave-aegean-sea-june-2021

Title	Short analysis of the wave field during the passage of IANOS medicane over the Ionian Sea
Date	5 October 2020
Activity	Type: Press release. Channel: Website
Audience	Type: Scientific community, Civil society. Size: 140
URL	https://poseidon.hcmr.gr/news/waves-medicane-ianos

Annexe 2.9 COSYNA (ID 17.1)

The following is the list of outreach activities reported by COSYNA, which belongs to Hereon.

Title	EuroGOOS FerryBox-HF Radar Joint Workshop
Date	17 March 2021
Activity	Type: Organisation of a workshop. Channel: Workshop
Audience	Type: Scientific community, Industry. Size: >100

Title	Information table at 2020 Ocean Sciences Meeting, San Diego, CA, USA
Date	16 February 2020
Activity	Type: Participation in a conference. Channel: Scientific Conference
Audience	Type: Scientific community. Size: 5000 participants
URL	https://www.agu.org/ocean-sciences-meeting

Title	Launch of the new version of visualisations for the European FerryBox database supported by COSYNA
Date	21 June 2021
Activity	Type: Website. Channel: Website
Audience	Type: Scientific community, General public. Size: >1000
URL	ferrydata.hzg.de

Annexe 2.10 HIDROGRAFICO+ (ID 19.1)

The following is the list of outreach activities reported by HIDROGRAFICO+, which belongs to IH.

Date	7 October 2020
Description	New to be published on JERICO website

Title	Participation with communication in the Webinar do Eurogeographics Knowledge Exchange Network
Date	8 July 2021
Activity	Type: Participation in a conference. Channel: YouTube
Audience	Type: Scientific community. Size: 50 participants
URL	https://www.youtube.com/watch?v=ni7Tcq6qNeE

Title	Participation with communication to the 6 Hydrographic Engineering Conference / 1 Portuguese-Spanish Hydrography Conference
Date	5 November 2020
Activity	Type: Participation in a conference. Channel: Website
Audience	Type: Scientific community. Size: 50-100
URL	https://jornadas.hidrografico.pt/index/en

Title	Participation with oral communication at the commemorative session of the World Hydrographic Day2021
Date	21 June 2021
Activity	Type: Participation in a conference. Channel: Website
Audience	Type: Scientific community, Industry, Policymakers. Size: 50
URL	https://www.hidrografico.pt/noticia/801

Annexe 2.11 VOS Finnmaid GHG - BGC (ID 21.1)

The following is the list of outreach activities reported by VOS Finnmaid GHG - BGC, which belongs to IOW.

Title	Daten des IOW belegen Rekord-Rückgang fossiler CO2 Emmissionen
Date	8 January 2021
Activity	Type: Press release. Channel: Newspaper
Audience	Type: Civil society, General public, Media. Size: circulation of 110000 copies

Title	Globale CO2 Bilanz 2020
Date	2 January 2021
Activity	Type: Press release. Channel: Email and newsletter, Newspaper
Audience	Type: Civil society, General public, Media, Customers. Size: 200 000 potential readers

Title	Kontinuierliche Messungen von Spuren- und Treibhausgasen auf der Fähre Finnmaid
Date	1 July 2021
Activity	Type: Website. Channel: Website
Audience	Type: Scientific community, Civil society, General public, Media, Customers. Size: to be added

Title	Meeting Finnmaid
Date	9 August 2021
Activity	Type: Participation in an event other than a conference or workshop. Channel: Twitter
Audience	Type: Scientific community, Customers. Size: 8 members

Annexe 2.12 NorFerry/NorSOOP (ID 25.1)

The following is the list of outreach activities reported by NorFerry/NorSOOP, which belongs to NIVA.

Title	UN Ocean Decade Laboratory - Scientists for Ocean Literacy
Date	8 July 2021
Activity	Type: Organization of a workshop and participation in a conference. Channel: Twitter, Presentation and panel discussion at workshop/conference
Audience	Type: Scientific community, General public, Policymakers. Size: 50
URL	https://eurogoos.eu/2021/07/01/scientists-for-ocean-literacy-eurogoos-activity-as-part-of-the-ocean-decade-inspiring-engaging-lab/

Annexe 2.13 PORTUS Observing and Forecasting System (ID 28.1)

The following is the list of outreach activities reported by PORTUS Observing and Forecasting System, which belongs to PdE.

Date	20 July 2020
Description	Publication: De Alfonso M, García-Valdecasas JM, Aznar R, Pérez-Gómez B, Rodríguez P, de los Santos FJ, Álvarez-Fanjul E. Record wave storm in the Gulf of Cadiz over the past 20 years and its impact on harbours. CMEMS OSR4, Chapter 4, Section 4.6., Journal of Operational Oceanography (In Press).

Date	1 July 2020
Description	New high-resolution models in PdE OpenDap for Gijón, Alicante, Cartagena, Avilés, Palma, Mahón, Ibiza

Title	Publication of UNESCO(IOC): update on sea level data QC/processing best practices, led by Puertos del Estado
Date	26 June 2020
Activity	Type: Press release. Channel: Website, Twitter
Audience	Type: General public. Size: General public, undetermined
URL	http://www.puertos.es/es-es/Paginas/Noticias/UNESCOPUERTOS26062020.aspx

Title	Keynote presentation EOF 2021: Harbour scale oceanographic processes and coastal feedbacks. By Enrique Álvarez Fanjul on behalf of Puertos del Estado team and collaborators
Date	5 May 2021
Activity	Type: Participation in a conference. Channel: Virtual conference: https://eof2020.es/, VI Expanding Ocean Frontiers conference (EOF 2021)
Audience	Type: Scientific community. Size: 100
URL	https://eof2020.es/the-meeting/programme

Title	Keynote presentation: "On the effect of sea level increases during the storm Gloria" (Pérez Gómez, B.)
Date	29 April 2021
Activity	Type: Participation in a conference. Channel: Virtual conference: https://eof2020.es/
Audience	Type: Scientific community. Size: 100
URL	https://meetingorganizer.copernicus.org/EGU21/session/39707#vPICO_presentations

Title	Twitter on news about PdE Cuadro de Mando Ambiental
Date	20 June 2021
Activity	Type: Social Media. Channel: Twitter
Audience	Type: Scientific community, Industry, Civil society, General public, Policymakers, Customers, Harbour and coastal engineering community. Size: 16000
URL	https://twitter.com/PuertosEstado/status/1406537153773965315

Title	Twitter on new PdE online service for historical data download
Date	25 May 2021
Activity	Type: Social Media. Channel: Twitter
Audience	Type: Scientific community, Industry, Civil society, General public, Policymakers, Media, Investors, Customers, Harbour and coastal engineering community. Size: 16000
URL	https://twitter.com/PuertosEstado/status/1397144238379438085?s=20

Title	Twitter on New Layer with animated particles for currents, wind and waves forecast maps
Date	13 March 2021
Activity	Type: Social Media. Channel: Twitter
Audience	Type: Scientific community, Industry, Civil society, General public, Policymakers, Media, Investors, Customers, Harbour and coastal

	engineering community. Size: 16000
URL	https://twitter.com/PuertosEstado/status/1370677282092777475

Title	Twitter on PdE iMar app for met-ocean information
Date	17 July 2021
Activity	Type: Social Media. Channel: Twitter
Audience	Type: Civil society, General public, Media, Customers. Size: 16000
URL	https://twitter.com/PuertosEstado/status/1416323048425279493

Title	Twitter and Linkedin post on PdE published articles about Gloria storm
Date	11 June 2021
Activity	Type: Social Media. Channel: Twitter, Linkedin
Audience	Type: Scientific community, Industry, Civil society, General public, Policymakers, Media, Investors, Customers, Harbour and coastal engineering community. Size: 16000 (Twitter followers) and 10000 (Linkedin followers)
URL	https://twitter.com/PuertosEstado/status/1403305622838382601

Annexe 2.14 Swedish Oceanographic Data Centre (ID 32.1)

The following is the list of outreach activities reported by Swedish Oceanographic Data Centre, which belongs to SMHI.

Title	Ferrybox Workshop
Date	17 March 2021
Activity	Type: Organisation of a workshop. Channel: Email and newsletter, Twitter
Audience	Type: Scientific community. Size: 40

Title	HFR Workshop
Date	17 March 2021
Activity	Type: Organisation of a workshop. Channel: Email and newsletter, Twitter
Audience	Type: Scientific community. Size: 40

Title	Swedish Ocean Decade Session
Date	9 June 2020
Activity	Type: Participation in a workshop. Channel: Email and newsletter, Twitter
Audience	Type: Scientific community, Industry, Civil society, General public, Policymakers. Size: 30

Title	Universeum
Date	21 May 2021
Activity	Type: Participation in an event other than a conference or workshop. Channel: Email and newsletter
Audience	Type: Scientific community, Civil society. Size: 15

Annexe 2.15 SOCIB Data Centre Multi-Platform Observatory (ID 33.1)

The following is the list of outreach activities reported by SOCIB Data Centre Multi-Platform Observatory, which belongs to SOCIB.

Title	Check it out 👇👇👇 #VirtualAccess #JERICOS3 #JERICORI
Date	10 December 2020
Activity	Type: Social Media. Channel: Twitter
Audience	Type: Scientific community. Size: 676
URL	https://twitter.com/socib_icts/status/1337001318624464896

Title	'Glider Toolbox': A toolbox for glider data management
Date	21 January 2021
Activity	Type: Website. Channel: Website
Audience	Type: Scientific community, General public. Size: N/A
URL	https://socib.es/index.php?seccion=detalle_noticia&id_noticia=449

Title	Use the #SOCIB #THREDDS #DataServer #free of charge
Date	9 July 2021
Activity	Type: Social Media. Channel: Twitter
Audience	Type: Scientific community, Industry. Size: 1961
URL	https://twitter.com/socib_icts/status/1413433145790173184

Title	Virtual Access to coastal ocean data enabled through the SOCIB Thredds Data Server
Date	13 July 2021
Activity	Type: Website. Channel: Website
Audience	Type: Scientific community, Industry. Size: 557

URL	https://www.jerico-ri.eu/2021/07/13/virtual-access-to-coastal-ocean-data-enabled-through-the-socib-thredds-data-server/

Annexe 2.16 JERICO RI e-Infrastructure (ID 33.2)

The following is the list of outreach activities reported by JERICO RI e-Infrastructure, which belongs to SOCIB.

Title	📰 We have gone along to meet Juan Gabriel Fernández, head of #SOCIB Data Centre working on #JERICOS3 #virtualaccess improvement 👉He shares the advancements on the provision & facilitation of centralized access to @JERICORI resources https://cutt.ly/gheahvE #ICTSNews
Date	20 November 2020
Activity	Type: Social Media. Channel: Twitter, Facebook
Audience	Type: Scientific community, General public. Size: 1262
URL	https://twitter.com/socib_icts/status/1329758137646190594, https://www.facebook.com/ICTSSOCIB/

Title	A unique entry point to facilitate virtual access to coastal ocean resources within the JERICO-S3 project
Date	20 November 2020
Activity	Type: Website. Channel: Website
Audience	Type: Scientific community, General public. Size: 92
URL	https://www.socib.es/index.php?seccion=detalle_noticia&id_noticia=441

Annexe 2.17 SYKE-ALG@LINE (ID 34.1)

The following is the list of outreach activities reported by SYKE-ALG@LINE, which belongs to SYKE.

Date	27 August 2020
Description	Weekly algae reviews for summer 2020 are available at https://www.syke.fi/en-US/Current/Algal_reviews, and the whole summer period is summarised in the press release of "Algae bloom monitoring". Though JERICO-S3 is not explicitly mentioned, data is partly collected using Algaline ferries and Utö Station (part of VA of SYKE and FMI). https://www.syke.fi/en-US/Current/Press_releases/Summary_of_algal_bloom_monitoring_JuneAu(58270)

Date	4 June 2020
Description	Release of Finnish marine data portal, including Algaline data https://www.syke.fi/en-US/Current/MarineFinlandfi__The_treasure_trove_of_F(57547)

Title	Cyanobacteria information from the Gulf of Finland
Date	1 June 2021
Activity	Type: Website. Channel: Website
Audience	Type: Scientific community. Size: 10
URL	https://swell.fmi.fi/hab-info/index.html

Title	Summary of algal bloom monitoring June-August 2021: The hot weather boosted the growth of cyanobacteria in the early summer – still the amount of cyanobacterial blooms was below average since mid-July
Date	30 August 2021
Activity	Type: Press release. Channel: Website
Audience	Type: Scientific community, Civil society, General public, Policymakers, Media. Size: 1000
URL	https://www.syke.fi/en-US/Current/Summary_of_algal_bloom_monitoring_JuneAu(61391)

Annexe 2.18 Keri Island Research Station (ID 35.1)

The following is the list of outreach activities reported by Keri Island Research Station, which belongs to TALTECH.

Title	Facebook post with video about Keri station
Date	19 January 2021
Activity	Type: Social Media. Channel: Facebook
Audience	Type: General public. Size: 100

Title	Social media post
Date	27 May 2021
Activity	Type: Social Media. Channel: Facebook
Audience	Type: Media. Size: 100

Annexe 2.19 OBPS-OTGA (ID 36.1)

The following is the list of outreach activities reported by OBPS-OTGA, which belongs to IODE of UNESCO-IOC.

Date	1 June 2020
Activity	Type: Website content and blogging
URL	https://www.oceanbestpractices.org/projects/

Date	12 April 2021
Activity	Type: Conference and webinar
URL	https://imdis.seadatanet.org/

Title	Evolving and Sustaining Ocean Best Practices IV: OBPS Community Workshop
Date	18 September 2020
Activity	Type: Organisation of a conference. Channel: Website, Email and newsletter, Twitter, YouTube
Audience	Type: Scientific community, Industry, Policymakers, Customers. Size: 500
URL	https://www.oceanbestpractices.org/community-engagement/workshops/workshop-iv-2020/

Title	Pearlman, J., Simpson, P., Karstensen, J., Buttigieg, P.L., Pearlman, F., Waldmann, C. and Hoerstmann, C. (2020) Improving Global and Regional Ocean Observing Through Best Practices and Standards. IEEE Oceanic Engineering Society Newsletter, June 2020, pp. 17-21.
Date	1 June 2020
Activity	Type: Non-scientific and non-peer-reviewed publication (popularised publication). Channel: Magazine
Audience	Type: Scientific community, Industry, Civil society, Customers. Size: 2000
URL	https://ieeeoes.org/publications/oes-beacon/

Title	EMODnet Open Conference 2021
Date	14 June 2021
Activity	Type: Participation in a conference. Channel: Website
Audience	Type: Scientific community, Customers. Size: 200
URL	https://emodnetconference2021.eu/

Title	IOC Assembly Webinar Ocean InfoHUB - Latin America & Caribbean
Date	8 June 2021
Activity	Type: Webinar. Channel: Webinar
Audience	Type: Scientific community, Customers. Size: 50

Annexe 2.20 VLIZ Marine Data Archive (ID 38.1)

The following is the list of outreach activities reported by VLIZ Marine Data Archive, which belongs to VLIZ.

Title	Open Belgium presentation of Marine Data Management
Date	6 March 2020
Activity	Type: Participation in a conference. Channel: Website, Email and newsletter, Twitter
Audience	Type: Scientific community. Size: 10
URL	https://2020.openbelgium.be/session/breaking-out-research-data-cycle-20-years-marine-data-management

Title	Presentation: "MDA & IMIS, data repositories as treasure chests for documented data"
Date	28 February 2020
Activity	Type: Organisation of a workshop. Channel: Website, Email and newsletter, Twitter
Audience	Type: Scientific community. Size: 33
URL	https://lifewatch.be/en/advancing-data-stewardship

Title	Presentation of MDA for Andromeda project
Date	17 September 2020
Activity	Type: Training. Channel: Project meeting
Audience	Type: Scientific community. Size: 15

Title	Presentation of MDA for Fish Intel project
Date	11 May 2021
Activity	Type: Training. Channel: Project meeting
Audience	Type: Scientific community. Size: 15

Title	EMODnet Biology course
Date	8 June 2020
Activity	Type: Training. Channel: Website
Audience	Type: Scientific community. Size: 50
URL	https://classroom.oceanteacher.org/enrol/index.php?id=430

Title	Adding the Marine Data Archive to FAIRsharing.org
Date	30 July 2020
Activity	Type: Website. Channel: Website
Audience	Type: Scientific community. Size: >10
URL	https://fairsharing.org/FAIRsharing.CjHLQw

Title	Presentation of MDA in the presentation"ARMS workflow: photographic and genetic data from Autonomous Reef Monitoring Structures (ARMS) to track NIS colonisation of European waters and monitor long-term changes of marine hard-bottom communities (scientific side)."
Date	20 May 2021
Activity	Type: Participation in a workshop. Channel: Website, Email and newsletter, Twitter, YouTube
Audience	Type: Scientific community. Size: 55
URL	https://www.youtube.com/watch?v=frcpBr6_few

Title	MDA listed as a recommended data repository for Nature Scientific Data
Date	30 July 2021
Activity	Type: Social Media. Channel: Twitter
Audience	Type: Scientific community. Size: 1857 views on Twitter
URL	https://twitter.com/PatriciaMariaMC/status/1421032517491216384