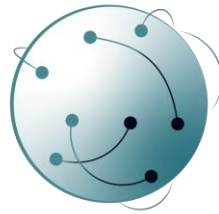




This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101059548.



**OPEN EARTH
MONITOR**

D2.3 Report “User requirements and data gaps”



Document control page

Project	Open-Earth-Monitor (OEMC)
Project, full title	A cyberinfrastructure to accelerate uptake of environmental information and help build user communities at European and global levels
Project number	101059548
Project start	June 1 st 2022
Deliverable	D2.3 Report "User requirements and data gaps"
Work Package	WP2 User-driven system design and FAIR workflows
Document title	D2.3 Report "User requirements and data gaps"
Version	1 st version
Responsible author	Nuno César de Sá (GFZ)
Contributors	Nora Meyer zu Erpen (GFZ), Martin Herold (GFZ)
Date of delivery	30th November 2021
Due date of deliverable	30th November 2021
Type	Report
Language	English
Status	<input type="checkbox"/> In progress <input checked="" type="checkbox"/> For review <input type="checkbox"/> Approved
Dissemination level	Public

Version history			
Version	Implemented by	Date	Description
V01	Nuno César de Sá (GFZ), Nora Meyer zu Erpen (GFZ), Martin Herold (GFZ)	30 Nov 2022	Final version



Contents

Document control page	2
List of Acronyms	4
Executive summary	5
1. Stakeholder engagement strategy	6
2. Methods for stakeholder need & requirement assessment	9
2.1 Broad survey on FAIR data	9
2.2 Use-case-related stakeholder interviews	10
3. Interim results of the broad survey on FAIR data	11
4. Conclusions	20
5. Next steps	21
6. Related tasks and outputs	22
Supplementary materials	23
A. Broad list of stakeholders	23
B. Broad survey on FAIR data – Question sheet	25
C. Draft interview guideline for use-case-related stakeholder interviews	28



List of Acronyms

AKIS	Agricultural Knowledge and Innovation Systems
CAP	Common Agricultural Policy
EC	European Commission
EO	Earth Observation
ESA	European Space Agency
EU	European Union
FAIR	Findable, Accessible, Interoperable, Reusable
GEO	Group on Earth Observation
GIS	Geographic Information System
IPCC	International Panel on Climate Change
JRC	The Joint Research Centre
NGO	Non-Governmental Organization
OEMC	Open Earth Monitor Cyberinfrastructure
SME	Small and Medium-sized Enterprises
UNCCD	United Nations Convention to Combat Desertification
WP	Work Package



Executive summary

The project Open Earth Monitor Cyberinfrastructure (OEMC) aims to maximize the impact and uptake of FAIR environmental data by collecting and analyzing EU and national level stakeholder needs and preferences. To achieve this goal, different activities are implemented in the framework of work package 2. This document gives an overview of current and future activities to push stakeholder engagement throughout the whole project duration.

The first part of the document describes shortly the overall stakeholder engagement strategy and builds on report D2.1. It gives an overview of different stakeholder groups and how their specific needs and feedback are collected for different aspects of the OEMC project.

The second part of the document describes the methods of the stakeholder need and requirement assessment based on an online survey and targeted interviews. The first interim results of the online survey will be presented and an interview guide for the targeted survey of use-case requirements for OEMC products and services will be presented.

This is the first version of the report on “User requirements and data gaps” showing the on-going and future activities. The outputs of the described activities are relevant for the tasks in WP3—6 as the stakeholder needs and feedback will have an impact on the design of particular functionalities of the OEMC computing engine (WP3), preparation and dissemination of the in-situ data (WP4), and especially on the EU and world monitors (WP5 and WP6).

A second version of this report will be published in December 2024 and will contain the final results of the online survey as well as the conducted outcome of the targeted interviews with respect to requirements and existing data gaps.



1. Stakeholder engagement strategy

The engagement of stakeholders is a key element of the OEMC project, as stakeholder needs and feedback will be systematically considered during the development of the OEMC product and services (WP3-6). The OEMC project aims to offer diverse services to heterogeneous stakeholders. These stakeholder groups can be roughly categorized thematically into four main groups, as shown in Table 1.

Table 1: A summary of the main stakeholder groups, categorized based on the type of OEMC services.

Stakeholder group	Main focus of services	Description
Government / public administration	CAP; AKIS; National spatial planning/agriculture/forestry agencies; urban and county-level planning; forestry agencies; risk assessment and disaster response agencies; spatial intelligence services	Significantly sized national or international institutions with impact goals at large and regional level with significant budget ability
Company / Industry	Geospatial industry and service providers; traffic/environmental management; Insurance sector such as Risk modeling, Loss assessment and Fraud detection; SME; EO and GIS start-ups	Institutions competing in the “private” sector, developing a diverse array of user-driven solutions.
Research and academia	Research institutes; Universities	Institutions which invest in the development of novel products or solutions, often with project-based or target-based funding.
NGO and citizens	Non-profit organizations; conservation and nature restoration organizations; open-source and open data users; private citizens	Large and diverse groups of users often operate voluntarily and under constrained budgets but can offer novel and smart solutions in often underfunded areas of application.

Stakeholder engagement, feedback and needs are acquired and considered in different ways, including:

- **Key stakeholder of the OEMC stakeholder committee:** This committee consists of (at least) eleven key stakeholders from relevant institutions (e.g., EC, JRC, UNCCD, IPCC, GEO). For the complete list please refer to Table 2 in report D2.1.



- **Use-case-related stakeholders:** These stakeholders are directly involved in each OEMC use-case and considered at the receiving end of the OEMC data streams of demonstration services. They have agreed to act as stakeholder and user with their expertise and interests in OEMC and will be engaged through regular interactions.
- **Broad geospatial community surveys:** Related efforts address all users, producers and providers of geospatial data willing to contribute to the OEMC project (e.g., via online surveys, open workshops). To address the challenge of identifying the general needs and requirements of a very diverse community, two broad groups are defined:
 - **Users** are individuals who primarily use geospatial and environmental data and products for their own tasks (e.g., decision making, research), but do not necessarily produce nor provide derived products.
 - **Producer & provider** are individuals who, on the one hand, use geospatial data and products (e.g., from the OEMC project), but at the same time are actively involved in the production and provision of geospatial and environmental products (e.g., maps, reports, statistics).

The following measures are or will be carried out within the framework of the project within the first project years:

- An **OEMC design workshop** was held between 18.07. - 20.07.2022 where a number of key stakeholders already participated. During this workshop an overall strategy to identify user needs was defined. For more information, please refer to report D2.1.
- The **OEMC stakeholder committee** is currently being established. The committee is composed of specific key stakeholders to provide general needs and feedback to the OEMC board. Furthermore, a continuous contribution and engagement of this group (e.g., contribution at conferences, joining panel discussions, keynote speeches at OEMC events) is expected throughout the project. For more information, please refer to report D2.1.
- Implementation of an **online survey to collect feedback on FAIR¹ data** from the broad EU and international geospatial community. The aim of this online survey is to get a comprehensive picture of whether users, producers and providers of geospatial and environmental data are aware of the FAIR data principles. In addition, it is investigated whether users and producers/providers have a similar or divergent understanding of the relevance of FAIR principles. This survey was started on 25.10.2022 and will remain open until December 2024. For this report, the results from 25.10.2022 to 16.11.2022 were evaluated and interim results are presented in chapter 3.

¹ Wilkinson, M. D., *et al* (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(1), 160018. <https://doi.org/10.1038/sdata.2016.18>



- Implementation of **targeted online interviews with key stakeholders** with focus on the OEMC use cases introduced by the 24 monitors (e.g., EU-coastal monitor, World-flood risk monitor). In a first step, use-case-related stakeholders will be interviewed with respect to their needs and requirements of the specific application.
- The preliminary list of potential use-case related stakeholders is shown in Table A of the supplementary materials. Through this direct interaction with stakeholders, a sustainable and long-term relationship can be established over the entire project duration. The content and structure of the interviews will be designed in cooperation with OEMC monitors representatives. A preliminary example is given in chapter 2 and section C of the supplementary materials.
- The **Open Earth Monitor Global Workshop²** will take place in Bolzano from 04.09. - 08.09.2023. This event aims to bring together European and global actors in the field of open-source earth observation applications in policy, business, research and for society. Registration and abstract submission are now possible.

The main activities to engage with stakeholders in the first 18 month of the project are shown in the Gantt chart:

Table 2: Planned activities on stakeholder engagement for the coming 18 months of the project.

Month	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Design Workshop																		
Establish Stakeholder Committee																		
Online Survey																		
Targeted Interviews (use cases)																		
Global Workshop																		

² The Open-Earth Monitor Global Workshop 2023 - <https://earthmonitor.org/gw2023/>



2. Methods for stakeholder need & requirement assessment

In addition to stakeholder events (e.g., workshops, conferences), different research and survey activities are implemented to collect and analyze stakeholder needs and requirements in the framework of WP2. Following activities will be described in the current chapter:

- A broad survey on FAIR data
- A targeted key stakeholder survey on needs and requirements on the OEMC monitors

2.1 Broad survey on FAIR data

The OEMC project aims to maximize the impact and uptake of FAIR environmental data. This online survey is a key activity to get a comprehensive picture of whether the broad geospatial community is aware of the FAIR data principles and what importance is attached to each principle. Furthermore, it is investigated whether users and producers & providers have a similar or divergent understanding of the relevance of FAIR principles³. To identify potential gaps, users and producers & providers are asked separate questions, the results of which can be compared. The survey consists of three question blocks:

- **The first block** includes seven general questions with respect to location of work, type of organization, role at work, main field of application, gender identity and range of age. In the last question, the participant must indicate whether she or he is primarily a user or producer/provider of geospatial data. According to the answer, the participant receives specific questions in the further course of the survey.
- **The second block** examines more information about the type of geospatial and environmental data that is primarily used or provided / produced.
 - Users are asked what type of geospatial data they primarily use, what properties of that data are particularly important to them and whether certain properties are problematic. The questions indirectly refer to the relevance of the FAIR principles (without the participants necessarily being aware of it).
 - Providers & producers are asked which type of data they offer, what they believe are the properties of geospatial data that are important and what properties are problematic for users.
- **The third block** focuses on the FAIR data principles. Users and producers & providers are asked whether they are familiar with the FAIR principles and which ones are particularly important to them. Users are also asked what they see as the biggest barriers to using FAIR and open data. While producers & providers are asked what primarily prevents them from offering fair and open data.

³ GOFAIR data principles: <https://www.go-fair.org/fair-principles/>



The survey⁴ includes 17 mandatory questions (and some optional response options) and was intentionally kept short so that many people from the geospatial community would respond to get a comprehensive picture. It will remain open for another 24 months to be completed by the next version of this report. In the meantime, the survey will be promoted at relevant events, conferences and through social media. This survey was created and published using EU Survey⁵ which is a tool developed by the EU commission for survey purposes and has been used in prior scientific studies on this field of earth observation⁶. The full sheet of questions of this survey is made available in Table B of the supplementary materials.

2.2 Use-case-related stakeholder interviews

As part of the project, a total of different OEMC monitoring tasks (WP 5 - 6) are to be developed and undertaken. These provide the technical underpinnings for the development of specific use-cases. These use-cases are characterized and driven by specific users/stakeholders that are actively involved in its definition, implementation and assessment. WP2 efforts will have an important role in terms of preparing and managing these stakeholder interactions in systematically acquiring user requirements, stimulating exchange among data producers and users and soliciting the user feedback once the data and information have been delivered.

Initial discussions with specific stakeholders have already taken place for some use-cases. However, in order to gain a comprehensive understanding of requirements for all monitors, targeted interviews will be conducted in the coming months with use-case related stakeholders, the OEMC monitor-leaders and WP2 representatives. This interaction aims to identify not just the needs of stakeholders but also to actively integrate them in the development of final products to be delivered by OEMC. A draft of the use case survey is provided in the Supplementary materials. The use-case survey will likely be changed and adapted according to each stakeholder and in coordination with the leading partners which are assigned.

The targeted interview consists mainly of three question blocks:

- The first block contains the **questions of the online survey on FAIR data** (see Table B in Suppl. Materials). In addition to the broad survey of the geospatial community, it is very important to understand to what extent use-case-related stakeholders are familiar with the FAIR data principles and what the individual principles mean to them. Their assessments can be of significant importance for the development of OEMC monitors.

⁴ Broad survey on FAIR data - <https://ec.europa.eu/eusurvey/runner/OpenEM-Survey-FAIR-geospatial-data>

⁵ EUSurvey - <https://ec.europa.eu/eusurvey/home/welcome>

⁶ Wagemann, J., Siemen, S., Seeger, B., & Bendix, J. (2021). Users of open Big Earth data – An analysis of the current state. *Computers & Geosciences*, 157, 104916. <https://doi.org/10.1016/j.cageo.2021.104916>



- The second block relates to **user-centered questions about the required products and information**. To identify the needs and requirements of the use-case-related stakeholders and their environment it is necessary to understand
 - Broad motivation and needs for specific use cases, incl. the field of application and related tasks in detail
 - existing routines and current work environment to fulfill these needs today; incl. Limitations and FAIR and open data status
 - currently used products/information and their characteristics to identify potential gaps in existing solutions
 - Expected information coming from OEMC use case and how and for what it should be used
 - type of users or services that will work with/use the product

After this block of questions, the required product type and its main properties must be known in order to define the specific product requirements in the next step.

- Based on the identified product type, a block of **detailed questions on the product requirements** follows. Due to the different monitors and diverse stakeholder groups, it is not possible to create a completely standardized questionnaire. The aim is to find suitable requirements to the following categories:
 - Content & units (e.g. definitions, parameters, indices, feature classes, objects)
 - Spatial level of detail (e.g. spatial resolution, minimum mapping units)
 - Spatial coverage (e.g. spatial extent, areas of interest)
 - Temporal detail (e.g. temporal resolution, temporal coverage, update frequency)
 - Accuracy (e.g. horizontal resolution, thematic accuracy, probability)
 - Access & delivery (e.g. via ftp server, data portals, web services, WMS & WFS)
 - Data formats (e.g. GeoTIFF, geojson, shapefile)
 - Metadata
 - Other requirements identified in the interview

The implementation of the interviews is planned for Q1/2023. It is planned to schedule 30 - 60 minutes per interview. In order to be able to focus on question blocks 2 and 3, the stakeholders will be asked to fill-in the questions of the online survey before the interview. If some stakeholders are not available for an interview, they will be asked to answer the questions in writing.

3. Interim results of the broad survey on FAIR data

In this chapter, the interim results of the online survey are presented with a focus on the results which we found to be more significant. While not all plots and tables are shown here for simplicity,



they are all made available in a purposely created GitHub repository⁷. Alongside, all the data used on the analysis, higher resolution images and the R scripts used are also available in that repository.

The results presented cover the period Oct. 25 - Nov. 14, 2022. As no events (e.g., workshops, conferences) were held during this period, the survey was distributed primarily through OEMC project partners email distribution lists and social networks. Thus, **a total of 114 complete responses** were acquired and are the basis for this first analysis.

The **first block of questions** consists of general questions to characterize the respondents (e.g. location of work, type of organization, role at work, main field of application, gender identity and range of age). In total, 114 people from 17 different countries participated (of which 102 from 14 EU countries). As Figure X shows, most participants work in Germany (54%, n=61), followed by Italy (10%, n=11), the Netherlands (8%, n= 9) and Romania (8%, n=9). For a better overview, all countries with two or fewer participants have been combined (Figure 1).

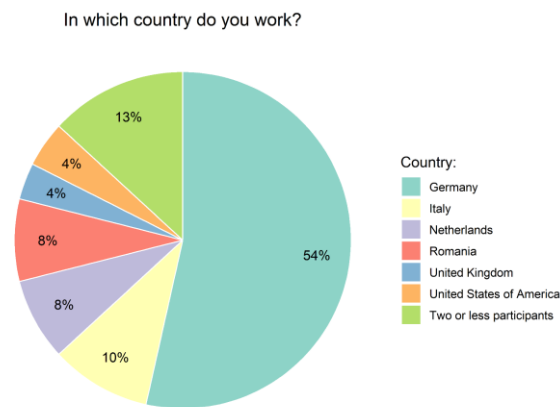


Figure 1: Summary of the survey answers regarding the workplace of the participants. A number of countries had 2 or less responded and were aggregated into a single table to facilitate visualization. The countries are Austria, France, Spain and Switzerland with only 2 participants and Canada, Croatia, Greece, Luxembourg, Portugal, Slovenia and Sweden with one participant each.

Most of the participants work in academia with 43% responding they work in a research institute plus 15 % in a university. In addition, many people from the governmental and public sector (19%) and the private sector (18%) participated. It is therefore not surprising that most of the participants are scientists (66%), followed by technicians (18%). Furthermore, a significant number of participants (13%) identified “other” as their job role, showcasing a strong diversity ranging from company founders and owners, GIS technician, system administrator, to sales and department heads. In terms of gender diversity, 65% of participants are male, 30% of participants are female and 2% of participants are non-binary / non-conforming. Furthermore, the majority of participants

⁷ Data, R scripts and higher quality plots are all available in the following repository: https://github.com/nunocesarsa/OpenEM_interimFAIRSurveyResults



were aged between 30 - 40 years (46%) with also a significant contribution of participants aged between 40 - 50 (27%) and 20 - 30 (17%).

In terms of field application, the results show a high diversity and given that respondents could answer multiple options, it is likely that many of them work in multiple fields. *Agriculture / land degradation* alongside *Nature conservation / biodiversity* were selected by 20% of the respondents and are the most common field of applications (Figure 2). Other fields like *Risk / hazards*, *Coastal marine areas* and *Water resources* also were selected a significant number of times.

What is your field of application?

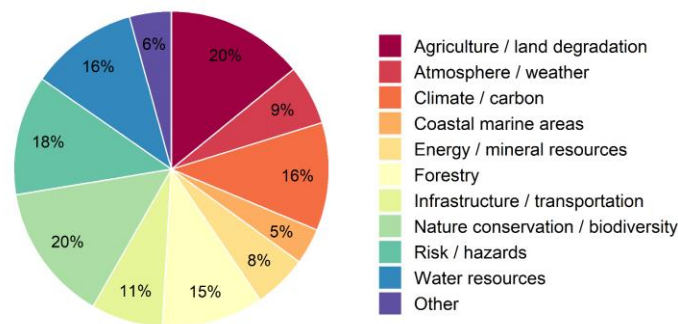


Figure 2: In this question, the participants could choose any option they wanted. The ratios represented on the pie chart show how many of the total participants (n=114) selected each of the options.

In the last question of the first block, participants must decide whether they are primarily users or providers of data. The result shows that most participants describe themselves as *users* (72%), while 28% are data *producers/providers* (Figure 3). Based on the answer, the participants were further given slightly different questions in the next two question blocks.

You are primarily a ...

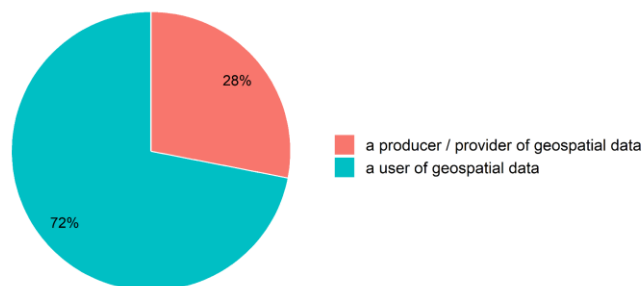


Figure 3: Overview of how respondents identified as part of the user group or the producer / provider group.

As mentioned before, the second block of questions refers to the characteristics of geospatial data as seen by its users on one side, and by producers/providers on the other side. The results of both groups are shown and compared in the further course. Figure 4 shows which data users



work with most often and which data is provided by producers/providers. The results show that *Open Satellite-based remote sensing data & derived products* are primarily used by many users (88%) and provided by most producers/providers (69%). The result indicates a high demand for open satellite-based earth observation data in the geospatial community.

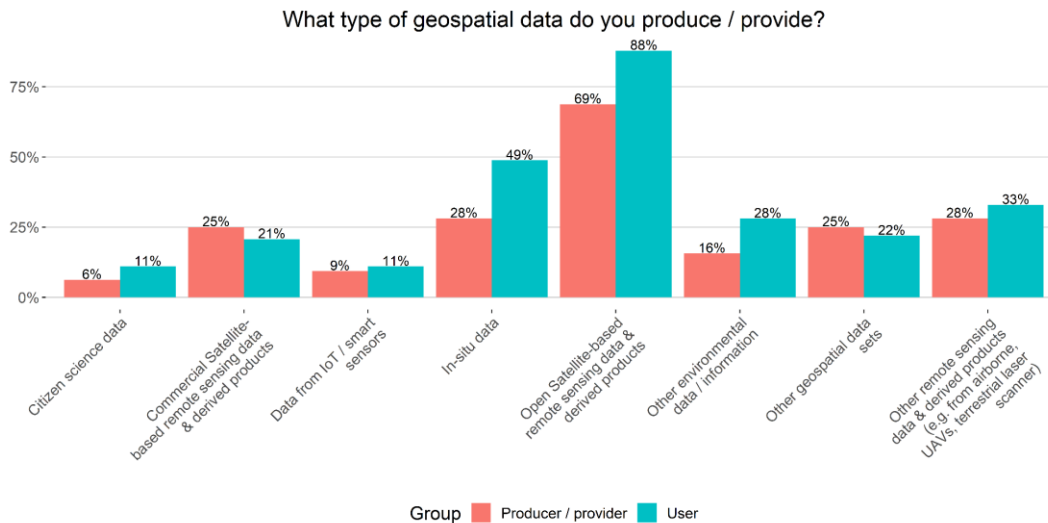


Figure 4: Comparison of responses between the group of users and producers/providers regarding the question on the type of geospatial data that is used or produced / provided

Furthermore, the scale level at which users and producers/providers operate was compared. While 56% producers/providers offer data products on a global scale, only 32% of users use global data. In total, 89% of the users work at local and federal scale, but only 56% of producers/providers make data available at that level of scale (Figure 5). This can indicate a gap between the user needs and data availability.

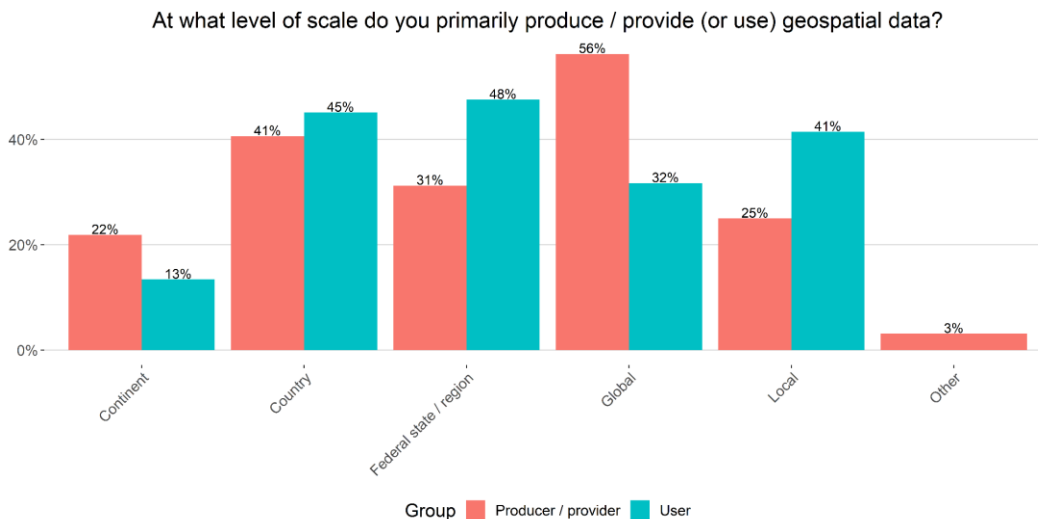


Figure 5: Comparison of responses between the group of users and producers/providers regarding the geographical scale at which they work on



Regarding important features of geospatial data, there is a general agreement between users and producers/providers (Figure 6). In particular, both groups find it extremely important that *Data is easy to find online* (57% of users and 69% of producers/providers) and that *Data is open* (54% of users and 59% of providers/producers). Less important features for both groups are 1) *Data can be retrieved via domain relevant community standards*, 2) *Data can be reproduced* and 3) *Data is interoperable with other data sets*. In the context of the FAIR principles, it shows that both groups prioritize easy findability and quick access to data instead of interoperability, reproducibility or clear community standards (Figure 6).

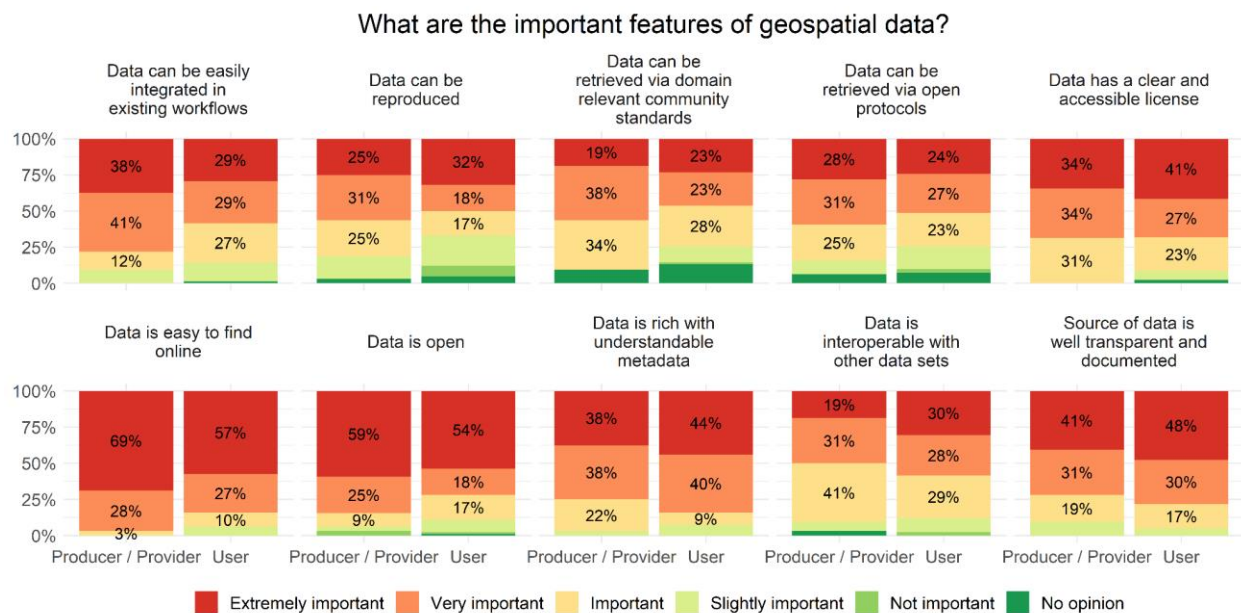


Figure 6: Overview of the most important features of geospatial data from the point of view of users and producers/providers of geospatial data.

In addition, users and producers/providers see similar problems regarding geospatial data (Figure 7). Data not being open is extremely problematic for many users (44%). Producers/providers also see this as a major problem for users. Another big problem for the users is incomplete metadata (for 20% of the users this is extremely problematic). Identification or registration to get access to data is not critical for most users. Nor do the producers/providers believe that this is a major problem for users.

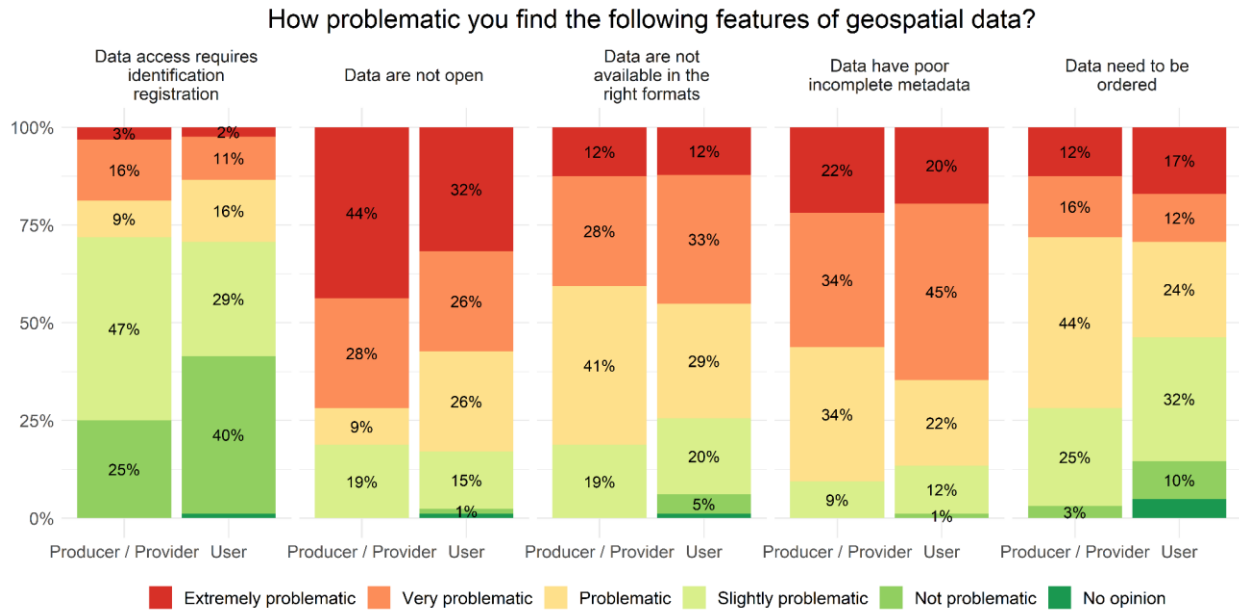


Figure 7: Overview of the most problematic features of geospatial data from the point of view of users and producers/providers of geospatial data.

The final question of this second block asked users what was their preferred method to find geospatial data and producers/providers how they delivered their data. Again, this question allowed respondents to provide multiple answers and therefore the ratios represent a total of times each answer was given by the total number of participants of that group, implying that they can sum to over 100%. Regarding producers/providers, 81% provide their data through a website and 59% through a geospatial catalog / geoportal. The preferred method for users to find geospatial data was through web searches (79%) or data hubs (e.g., Sentinel hub) (55%) or geospatial catalogs (50%). Overall, it appears a bit of mismatch exists between both groups with providers focusing on providing data through a “website” instead of data hubs which users seem to prefer.

The third block of questions relates especially to the FAIR data principles. In this case, we found that a significant proportion of both users (35%) and producers/providers were not familiar with FAIR data principles (28%), as shown in Figure 8. Nevertheless, while 56% of producers/providers are (at least) familiar with the FAIR principles, only 36% of the users are familiar with the FAIR principles. These results seem to indicate that in the case of our participants, there is a relatively poor familiarity with FAIR data principles for both groups, even if a bit less in the case of producers/providers. For more details on each of the results please refer to GitHub repository mentioned before.

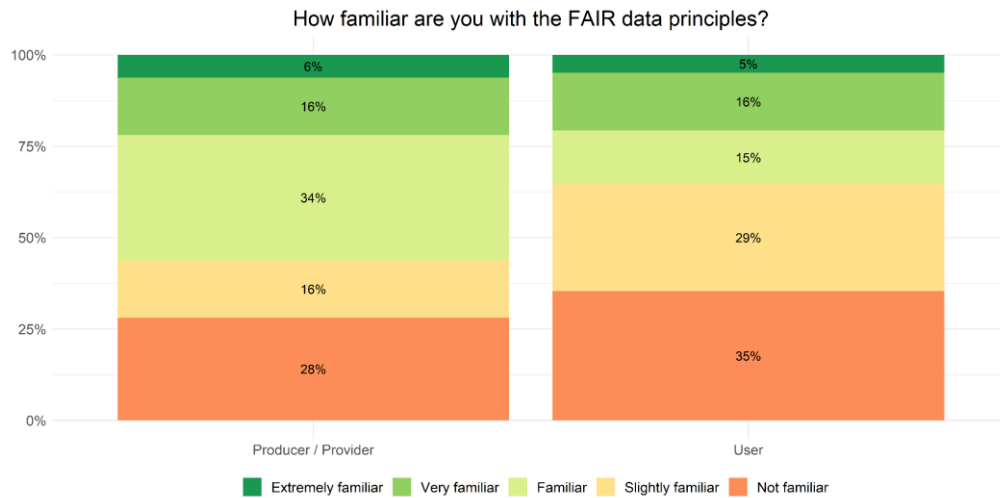


Figure 8: Overview familiarity of both producers/providers and users regarding the FAIR data principles.

When asked which FAIR principles are particularly important to users and producers/providers, it is again apparent that the priority of both groups is on *Data must be easily accessible* and *Data must be easy and quick to find* (Figure 9). Again, similarly to a previous result (Figure 6 and Figure 7), less importance is given to reproducibility of data or interoperability with other data.

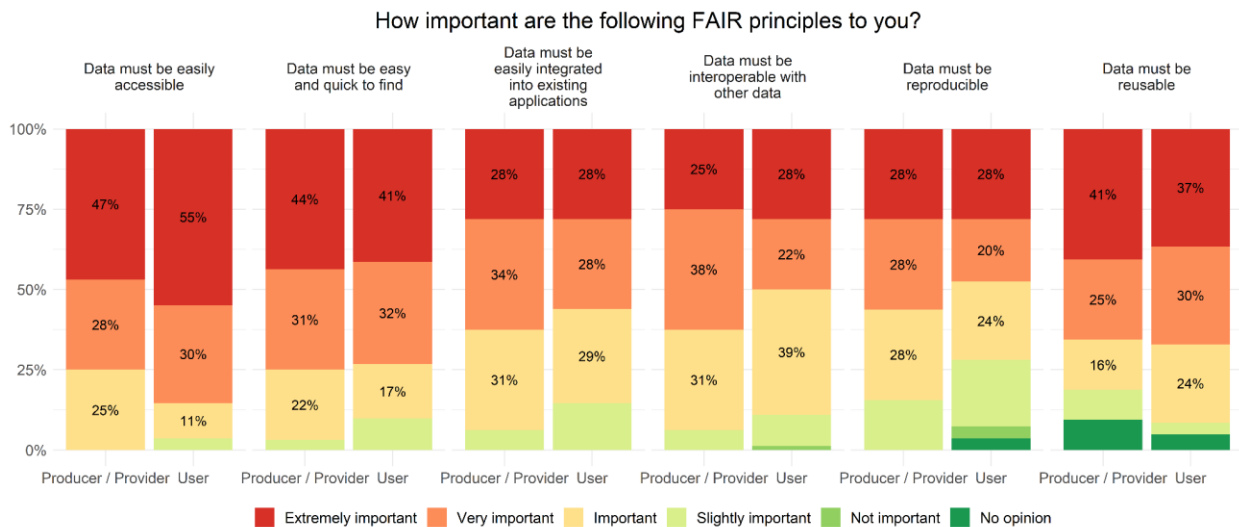


Figure 9: Overview importance that both producers/providers and users give to the core FAIR data principles

A very significant proportion of producers/providers responded that they are already providing their data at least partially according to the FAIR principles with 44% responding that they provide FAIR data and 31% saying that they provide partially FAIR data (Figure 10). On the other side, only 18% of the users claimed to have used FAIR data while 24% claimed to have not used fair data (Figure 10). The large majority of users do not know if they are using FAIR data or not (57%). These results represent a large lack of awareness from the users regarding the source of their



data which can obviously be problematic for providers/producers and their concerns regarding data sharing.

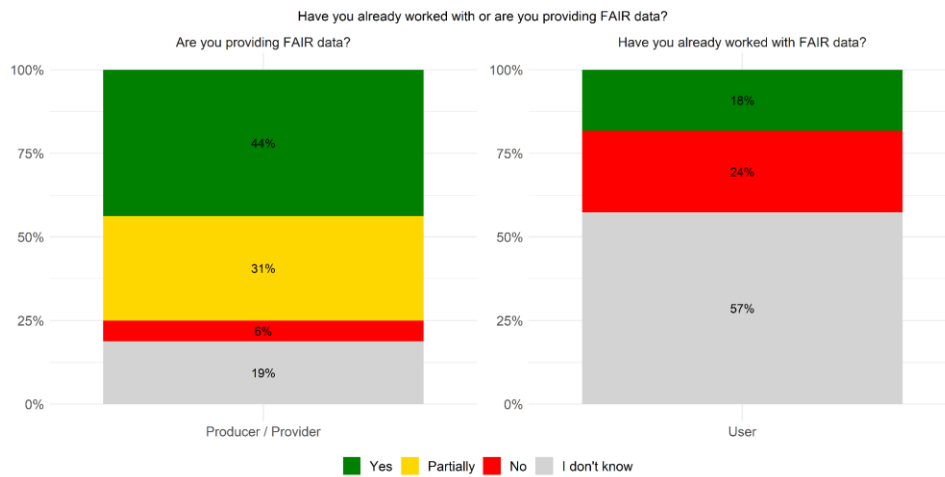


Figure 10: Overview of the responses to the question if users have worked with FAIR data and if producers/providers already provide data according to FAIR standards.

Overall, most producers/providers (56%) correctly identified that FAIR data does not necessarily is open. In comparison, 38% of the users (38%) expected FAIR data to be open (Figure 11). This again shows that the producers/providers are more familiar with the FAIR principles. Also, of significant note is that the majority of users does not know the difference between FAIR and Open data (48%). This might indicate a need for addressing this in future communications with the broader geospatial community.

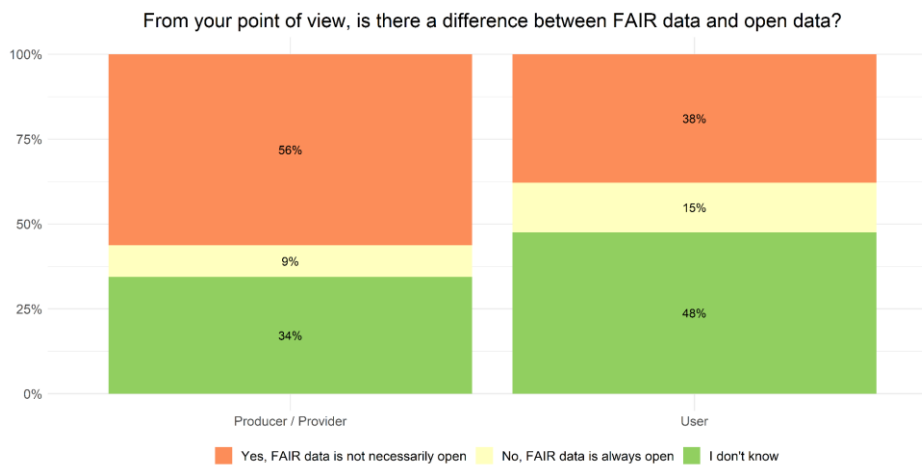


Figure 11: Comparison of the expectation between users and producers/providers regarding FAIR data being also Open data.

Furthermore, producers/providers were asked what they perceive as barriers to produce or provide FAIR data (Figure 12). The main reasons cited by producers/providers are *lack of resources* (69%) and *missing incentives* (50%). A significant proportion also pointed out a *lack of guidelines* and *potential misuse* (34% each). Interestingly, the *lack of technical solutions* or the *risk of competitive disadvantages* are not considered as significant barriers. This may indicate

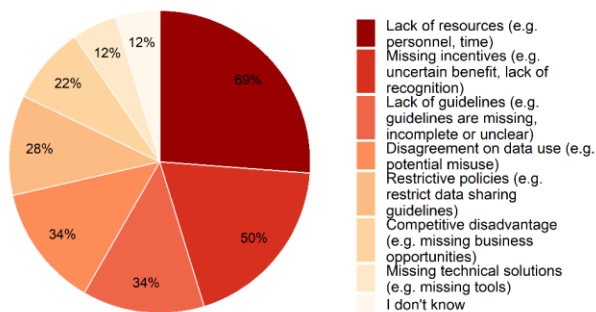


that producers/providers are potentially ready to offer FAIR data but do lack the incentive to do so.

In addition, users were asked what they perceive as barriers to using FAIR data. From users' perspective, the "lack of awareness of which data is FAIR" (49%) and the *lack of knowledge of the benefits* (43%) of FAIR data are the most important barriers for FAIR data use. On the other side, *missing technical solutions* (18%) or *concerns about licensing* (20%) are the least significant barriers from the perspective of this group. The results of both groups show that awareness and incentives are critical blockers to use and produce FAIR data. Technical options or licensing solutions are perceived as rather uncritical by both groups.

Barriers to the use of FAIR data from:

a producers / providers perspective



a user perspective

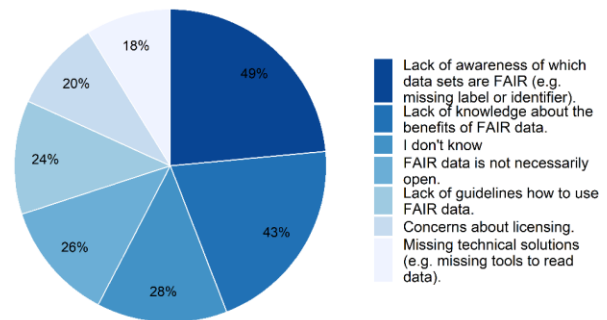


Figure 12: Overview of the most significant barriers for users to use FAIR data and for producers/providers to offer FAIR data. The respondents could choose up to three options and therefore the % represent the total number an answer was given by the total number of participants of that group

Regarding Open data, *Competitive disadvantage*, *Economic disadvantage* and *Concerns regarding sensible data* were the most selected options with 53%, 50% and 47% of the producers/providers group (Figure 13). This clearly shows that this group is mainly concerned with competition, economic benefit and data protection as indeed these aspects play a significant role in the valuation of products. On the other side, concern about privacy, ethical use of data or restrictive policies are not considered to be significant barriers which is not surprising considering that the geospatial data field often works with broad and aggregated data which inherently protects many of these aspects.

From the user perspective, responses are more well spread throughout all the options with the main barriers of Open data being the *lack of continuity* (50%), *missing support* (44%) and *lack of standards* (43%) (Figure 13). Still, *metadata concerns* (40%) and *licensing* (38%) were highly voted options with the least voted being technical difficulties (20%) (Figure 13). Overall, users are concerned that Open data might not be consistently available in time and there will be a lack of support and standards which are barriers that can be potentially addressed with Open and FAIR data.



Barriers to the use of Open data from:

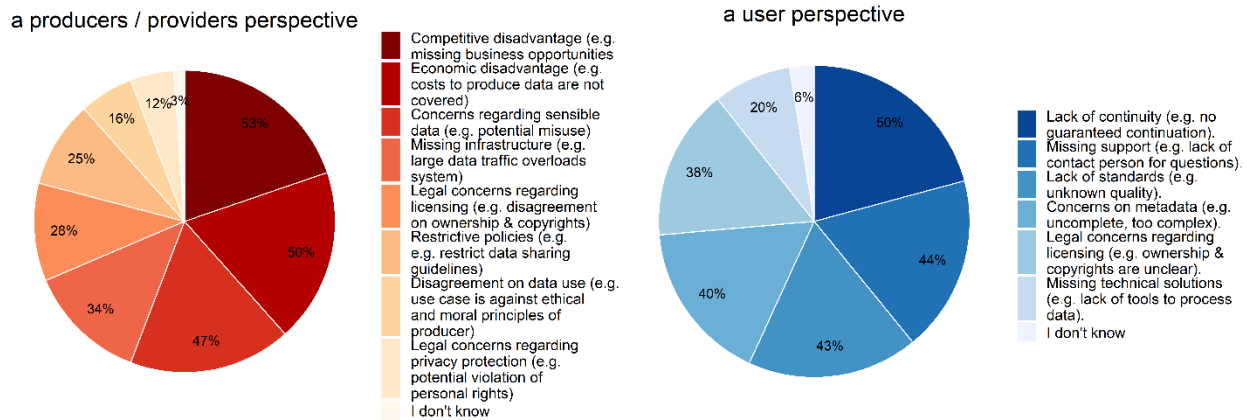


Figure 13: Overview of the most significant barriers for users to use Open data and for producers/providers to offer FAIR data. The respondents could choose up to three options and therefore the % represent the total number an answer was given by the total number of participants of that group

4. Conclusions

Within the framework of the OMEC, a comprehensive stakeholder engagement framework is implemented to systematically identify stakeholder needs and existing data gaps. For this purpose, various activities are carried out, such as the establishment of a stakeholder committee, the implementation of workshops and conferences, as well as an online survey and use-case-related interviews.

In this report the first intermediate results of the online survey on FAIR data are published. Within the first three weeks, 114 people took part in the survey, which is a good participation rate. The survey was widely disseminated and promoted, particularly in the German research community, resulting in biased results in terms of geographic distribution (54% of the participants work in Germany) and stakeholder types (58% of participants work in academia). It has also been shown that so far it has been primarily men (65%) who have taken part in the survey, while women (30%) or other gender identities (2%) are underrepresented. There is a great variety in the field of application, which shows that many different fields of work of the geospatial community were reached. To reach even more participants from different European countries, the survey will remain online until the second version of this report (D2.9) in December 2024 and will be heavily promoted in the meantime at the OMEC website, social media (e.g. LinkedIn, Twitter), networks (e.g. female networks in the geospatial community) and relevant events (e.g. EuroGEO⁸, Open Earth Monitor Workshop 2023).

⁸ EuroGEO Annual meeting to be held in April 2023 - <https://www.eurogeography.eu/conferences/>



In the context of geospatial data, it is particularly important for many users that the data is freely available and easy to find. In addition, great value is seen in complete metadata. With respect to the FAIR data principles, the interim results show that most respondents (especially users) have a very limited knowledge of the characteristics, availability and benefits of FAIR geospatial data.

This is also reflected in the fact that the main barriers to use FAIR data are *Lack of awareness of which data sets are FAIR* and *"Lack of knowledge about the benefits of FAIR data"*. Also, the fact that *FAIR data is not necessarily open* is an issue for many users concerning FAIR data and something that needs to be clarified. Many respondents do not necessarily expect FAIR data to be open, however, for many users having open data is a high priority. The survey results point at some of the current gaps and limitations for applying the FAIR principles for environmental data. FAIR principles must be made known and promoted as part of the OEMC project. This needs to be a focus in direct interaction with stakeholders (e.g., during interviews, on workshops) and a sustained dialog between data producers/providers and users (in particular in the context of the OEMC use cases).

5. Next steps

Based on the concepts and preliminary results for assessing user requirements and gaps in this deliverable, the following main next steps are foreseen with respect to that part of the project work:

1. Expand the survey for the broader geospatial community to reach a more diverse audience (target specific communities) and consolidate and disaggregate the findings from the survey for the next version of the deliverable (D2.9: Report "User requirements and data gaps" final version). Dedicated promotion activities are planned.
2. Implement the systematic stakeholder engagement around the OEMC use cases. The concept and approach is presented in this deliverable and implementation will start in 2023.
3. Refine the analysis of user requirements and gaps for the next version of the deliverable (report D2.9). The main inputs are coming from the broad stakeholder survey, the OEMC use case interactions and developments, and other stakeholder engagement efforts from the project (i.e., Open Earth Monitor Workshop 2023)
4. Develop and implement a user feedback mechanism. The concept will be presented in the next version of the deliverable (report D2.9) and will form the underpinnings for report D2.12: "Assessment of system usability" to be prepared towards the end of the project.



6. Related tasks and outputs

Report D2.1: Stakeholder Committee and "Open-Earth-Monitor design" workshop.

Report D2.2 (1st version): Status and prospect for European environmental data

Report D2.4 (1st version): Economic Assessment Framework Guidelines



Supplementary materials

A. Broad list of stakeholders

Table of the broad list of stakeholders that have been currently identified of potential interest for the Open Earth Monitor project. Some of these stakeholders have already been contacted while others are to be contacted. It is likely that this list will change in the future as new stakeholders are identified and contacted. Also, these stakeholders might be in contact with multiple partners within the project depending on the different use-cases that have been developed.

Table A – List of potential use-case stakeholders that have been or will be invited to participate in Open Earth Monitor

Stakeholder	Leading partner(s)	Domain	Target user communities
AdbPo	CNR	Public sector	
Agroseguros S.A.	GILAB	Industry user	Agricultural insurance companies
CSIRO	OGH	Research and Academia	UN organizations; National agencies; Reforestation/tree planting companies
DestinE	MPG	Public sector	
DPC	CNR	Public sector	
EEA	IFGI, OGH	Public sector	Health organizations; Research organizations & universities; National agencies; European Commission agencies; EU citizens; Regenerative Agri project developers
EFINET	OGH, WU	Public sector	Reforestation/tree planting companies; Forestry organizations; National agencies; NGOs; EU citizens
EPTB	CNR		
ESA	CNR, MPG	Public	
EuroGEO	OGH	NGO	Reforestation/tree planting companies; Forestry organizations;
ForestSat	GILAB	NGO	Forestry organizations
GEOM	OGH		Hydro-meteorology offices; UN organizations; Reforestation/tree planting companies



IDH	OGH	Industry user	NGOs; Reforestation/tree planting companies;
IITA	OGH	NGO	
INPE	WU	Public sector	Brazilian agencies e.g., INPE; Int not-for-profit WRI TNC CI;
JRC	CNR, MPG, OGH, WU	Public sector	European Commission agencies; EU citizens; Regenerative Agri project developers; National agencies; NGOs;
KARAVIAS INSURANCE S.A.	GILAB	Industry user	Agricultural insurance companies;
LAPIG/UFG	OGH	Research and Academia	Brazilian agencies e.g., INPE; Regenerative Agri project developers; Int not-for-profit WRI TNC CI
LDN-UNCCD	OGH	Public sector	UN organizations; National agencies; Reforestation/tree planting companies; NGOs; Int not-for-profit WRI TNC CI; Hydro-meteorology offices;
Mir	OGH	Industry user	NGOs; Reforestation/tree planting companies;
OCW	GILAB		Regenerative Agri project developers;
OECD/countries/GFOI	GFZ	Public sector	IPCC GPG/SEEA; UNFCCC;
Province of South Tyrol	EURAC	Public sector	Hydro-meteorology offices; Weather Monitoring agencies;
RER	OGH	Public sector	Health organizations; Research organizations & universities; SME's & startups
TNC	OGH	NGO	NGOs; Int not-for-profit WRI TNC CI;
UN Spider	TS	Public sector	
UNDP-HR	OGH	Public sector	National agencies; Forestry organizations;
WEF	MPG	NGO	
WMO	CNR	Public sector	
WRI	GFZ, OGH, WU	NGO	IPCC GPG/SEEA; UNFCCC; Brazilian agencies e.g., INPE; Regenerative Agri project developers; Int not-for-profit WRI TNC CI;



B. Broad survey on FAIR data – Question sheet

Table B: This table shows all the questions that were asked in the broad survey organized per each block as explained in the D2.3 Report. Optional questions are numbered with o1 or o2.

Block	Topic:	Question nr:	Target:	Question:
1	General information	1	Both	In which country do you work?
1	General information	2	Both	What type of organization do you work for?
1	General information	3	Both	Which role applies primarily to you?
1	General information	4	Both	What is your gender identity?
1	General information	5	Both	What are your main fields of application?
1	General information	6	Both	What is your Age
1	General information	7	Both	You are primarily... a) Answer a user of geospatial data or b) a producer / provider of geospatial data
2	Geospatial data characteristics	8	Producers / providers	What type of geospatial data do you produce / provide?
2	Geospatial data characteristics	8o1	Producers / providers	What kind of other environmental data sets do you produce / provide? (optional)
2	Geospatial data characteristics	8o2	Producers / providers	What other geospatial data sets do you produce / provide? (optional)
2	Geospatial data characteristics	8	Users	What type of geospatial data do you primarily use?
2	Geospatial data characteristics	8o1	Users	What kind of other environmental data sets do you use? (optional)
2	Geospatial data characteristics	8o2	Users	What other geospatial data sets do you use? (optional)
2	Geospatial data characteristics	9	Producers / providers	At what level of scale do you primarily produce / provide geospatial data?
2	Geospatial data characteristics	9o1	Producers / providers	At what level of scale, you primarily produce / provide geospatial data? (optional)
2	Geospatial data characteristics	9	Users	At what level of scale do you primarily work with geospatial data?



2	Geospatial data characteristics	9o1	Users	What level of scale you primarily work with geospatial data? (optional)
2	Geospatial data characteristics	10	Producers / providers	From your point of view: what are important features of geospatial data for the users?
2	Geospatial data characteristics	10o1	Producers / providers	Other features are extremely important to the users? (optional)
2	Geospatial data characteristics	10	Users	How important are the following features of geospatial data to you?
2	Geospatial data characteristics	10o1	Users	Other features that are very or extremely important to you? (optional)
2	Geospatial data characteristics	11	Producers / providers	From your point of view: how problematic do users find the following features of geospatial data?
2	Geospatial data characteristics	11o1	Producers / providers	Other features that are very or extremely problematic to users? (optional)
2	Geospatial data characteristics	11	Users	How problematic do you find the following features of geospatial data?
2	Geospatial data characteristics	11o1	Users	Other features that are very or extremely problematic to you? (optional)
2	Geospatial data characteristics	12	Producers / providers	How can users find the geospatial data you produce / provide?
2	Geospatial data characteristics	12o1	Producers / providers	Which other way do users find the geospatial data you produce / provide? (optional)
2	Geospatial data characteristics	12	Users	What is your favorite approach to finding geospatial data?
2	Geospatial data characteristics	12o1	Users	Which other approach you use to find geospatial data? (optional)
3	FAIR data	13	Producers / providers	How familiar are you with the FAIR data principles?
3	FAIR data	13	Users	How familiar are you with the FAIR data principles?
3	FAIR data	14	Producers / providers	How important are the following FAIR principles to you?
3	FAIR data	14	Users	How important are the following FAIR principles to you?
3	FAIR data	15	Producers / providers	Are you providing FAIR data?



3	FAIR data	15	Users	Have you already worked with FAIR data?
3	FAIR data	16	Producers / providers	From your point of view is there a difference between FAIR data and open data?
3	FAIR data	16	Users	From your point of view is there a difference between FAIR data and open data?
3	FAIR data	17	Producers / providers	What do you think are the biggest barriers to produce more FAIR data? Select your top 3!
3	FAIR data	17o1	Producers / providers	Any other barrier(s) to produce more FAIR data? (optional)
3	FAIR data	17	Users	What do you think are the biggest barriers to use FAIR data? Select your top 3
3	FAIR data	17o1	Users	Any other barriers to use FAIR data? (optional)
3	FAIR data	18	Producers / providers	What do you think are the biggest barriers to produce more open data? Select your top 3!
3	FAIR data	18o1	Producers / providers	Any other barrier(s) to produce more open data? (optional)
3	FAIR data	18	Users	What do you think are the biggest barriers to use open data? Select your top 3
3	FAIR data	18o1	Users	Any other barriers to use open data? (optional)



C. Draft interview guideline for use-case-related stakeholder interviews

Draft interview guideline for use-case-related stakeholder interviews

- **Questions of the online survey on FAIR data** (see Table B in Supplementary materials)
- **User-centered questions about the required product**
 - What is your field of application?
 - What are your specific use cases?
 - What is your main task? (e.g. reporting, analytics, research)
 - How are these tasks performed today?
 - What data or products are used?
 - Which are the working routines?
 - What works well?
 - Where do problems exist?
 - Which type of information is required? (e.g., quantity of biomass, areas prone to be flooded, forest clear cuts)
 - Which type of product is required? (e.g., analysis-ready data, maps, alerts, reports, statistics)
 - How will the product be used?
 - As a final product? (No further processing or adaption)
 - As an intermediate product? (e.g., further processing, implementation in existing workflows, adjustment of parameters)
 - Who will work with the product? (e.g., management, technician, scientist)
 - How technophile are users?
 - Which kind of (previous) knowledge do they have?
 - Where is the product used?
 - Any constraints on devices? (e.g., on standard PCs, mobile devices)
 - Any constraints on locations? (e.g., areas with unstable internet connection, in the field on mobile devices?)
 - How often will the product be used? (e.g., daily, weekly, annual)
 - Does the product have to be comparable to former products? (heritage)
 - Can you think of other stakeholders for whom this product is highly relevant?



- **Specific questions about the required product with respect to**
 - Are there **specific measurements or parameters** that should be used? (e.g., parameters, indices)
 - Do the results have to be in a **certain unit**? (e.g., pixel, square meters, volume estimation, feature classes)
 - What **level of spatial detail** must the data have? (e.g., spatial resolution, minimum mapping units)
 - What **level of scale** is required? (e.g., global, continental, local)
 - What is the desired **spatial coverage**? (e.g., spatial extent, areas of interest)
 - What is the desired **temporal coverage**? (e.g., temporal extend, update frequency)
 - Which **accuracy measurements** are important? Are there certain accuracies that must be achieved? (e.g., horizontal & vertical resolution, thematic allocation, probabilities)
 - How should the product be made **accessible**? (e.g., download via ftp server, data portals, web services, WMS & WFS, APIs)
 - Which **formats** are required? (e.g., GeoTIFF, geojson, shapefile)
 - What is the required **reference system**? (e.g., Lat-Long, projections)
 - Are there any specific requirements for the **metadata**?
 - Which **license** requirements must be met?
 - Are there any specific requirements for the **documentation** of the product?
 - Other requirements identified in the interview