Scientific Lake

Deliverable 6.1: Communication, engagement, and exploitation Plan

Due Date of Deliverable	30/06/2023
Actual Submission Date	30/06/2023
Work Package	WP6
Tasks	T6.1, T6.2, T6.3, T6.4
Type	R – Document, report
Approval Status	Accepted
Version	1.0
Number of Pages	39
The information in this document reflects only the author's views and the European Commission is not liable for any use that may be made of the information contained therein. The information in this document is provided "as is" without guarantee or warranty of any kind, express or implied, including but not limited to the fitness of the information for a particular purpose. The user thereof uses the information at his/ her sole risk and liability.	

Revision history

Version	Date	Reason	Revised by
0.0	25/05/2023	First Draft	Stefania Amodeo, Zenia Xenou, Natalia Manola
0.1	01/06/2023	Agreement on structure & References	Stefania Amodeo, Zenia Xenou, Thanasis Vergoulis
0.2	02/06/2023	Intermediate version	Stefania Amodeo, Zenia Xenou
0.3	15/06/2023	Peer review	Miriam Baglioni, Serafeim Chatzopoulos
0.4	29/06/2023	Peer review comments addressed	Stefania Amodeo, Zenia Xenou
1.0	30/06/2023	Final Version after proofreading	Stefania Amodeo

Author List

Organisation	Name	Contact Information
OpenAIRE	Manola Natalia	natalia.manola@openaire.eu
OpenAIRE	Xenou Zacharoula (Zenia)	zenia.xenou@openaire.eu
OpenAIRE	Amodeo Stefania	stefania.amodeo@openaire.eu

Contributor List

Organisation	Name	Contact Information
ARC-OpenAIRE	Thanasis Vergoulis	vergoulis@athenarc.gr
ARC	Serafeim Chatzopoulos	schatz@athenarc.gr
CNR	Miriam Baglioni	miriam.baglioni@isti.cnr.it

Table of Contents

List of Tables

List of Figures

Abbreviation List

AI - Artificial Intelligence

API - Application Programming Interface

SKGs - Scientific Knowledge Graphs

COARA - COalition for Advancing Research Assessment

EC - European Commission

ELG - European Language Grid

EOSC - European Open Science Cloud

ESFRI - European Strategy Forum on Research Infrastructures

NLP - Natural Language Processing

NOAD - National Open Access Desk

OS - Open Science

RDA - Research Data Alliance

SMEs - Small and Medium Enterprises

- **Disclaimer**

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

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

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

This publication has been produced with the assistance of the European Union. The content of this publication is the sole responsibility of the SciLake consortium and can in no way be taken as a reflection of the views of the European Union.

SciLake is a project funded by the European Union (Grant Agreement No 101058573).

- Executive Summary

This document presents the strategy and tools used to implement the SciLake Communication, Engagement and Dissemination plan. This plan designs an approach that promotes the project and its results to targeted audiences and all major stakeholders. It also lists all the tools and materials used to accomplish this task.

Recognising the importance of engaging with, and reaching out to, the target audience early on, this document outlines the definition of the SciLake core messages for each of the key stakeholder groups. It selects the most appropriate dissemination channels for the messages to be communicated and the key stakeholder groups to be reached.

Working towards a long-term impact for the project, the engagement of the major target stakeholder groups is fostered by providing a communication platform that documents project outcomes with success stories, policy briefings, methods, tools, and data.

This plan also contains the definition of activities that will enforce and spread the project's messages and results. It also includes a time plan that will serve as a guide for their delivery. It presents useful dissemination tools and materials that will be developed and used for the successful outreach and engagement of key target audiences.

Regular internal progress reports on the activities planned will provide a monitoring basis and inform any updates to the plan regarding the success of project results dissemination and the engagement of the key stakeholder groups.

1 Introduction

This deliverable describes the strategic plan for implementing the SciLake Horizon Europe project's Communication, Engagement, and Exploitation. The objective is to develop a strategy that maximises its impact and disseminates its results across multiple audiences.

This report also includes a description of all activities to engage the relevant communities in a continuous cycle, ensuring we elicit requirements and promote project results for adoption. It introduces valuable services and tools to be developed and implemented to achieve this goal.

Our methodology employs a phased and multidimensional approach based on the following elements:

- Identifying the key target groups, as well as future users.
- Engaging the targeted stakeholder groups.
- Identifying the best tools and practices to promote the uptake of SciLake's outputs.
- Communicating with the target audience using the appropriate channels and marketing materials.

The Communication, Engagement, and Exploitation plan will be regularly updated with the planned activities and will provide ongoing updates on the success of project results dissemination and the engagement of key stakeholders.

2 Objectives

The main objectives of SciLake to be promoted through the communication, engagement, and exploitation strategy are to:

- Introduce and establish the concept of the scientific lake, a research ecosystem where scientific knowledge is contextualised, connected, interoperable, and accessible overcoming challenges related to the heterogeneity and interconnectivity of the underlying data.
- Raise awareness and promotion of Scholarly and Scientific Knowledge Graphs (SKGs) as a core means for discovery, tracking, and monitoring of research.
- Establish a community that follows and drives (where possible) developments towards open, transparent, and interoperable SKGs (access, participation, governance).
- Improve the content coverage and quality of SKGs, promoting and facilitating the inclusion of heterogeneous content (e.g., including different types of research artefacts), developing measures to enhance the significance and impact of the content, and building multilingual bridges.

- Facilitate the creation, enrichment, management, and maintenance of different purpose SKGs and build upon and promote existing SKG interoperability practices and specifications towards a global network of interchangeable SKGs, catalysing the creation of domain-specific added-value services.
- Deliver added-value services on top of the SKGs demonstrating that the scientific lake concept can significantly reduce the respective implementation efforts.

3 Target groups and future users

Understanding all involved stakeholder groups from the outset is crucial for the success of SciLake. By investigating and comprehending the objectives and needs of each group, it will be possible to open communication channels with all parties involved. This will ensure their participation and dedication throughout the project, ensuring the achievement and sustainability of SciLake's results. Table 1 lists and describes our target groups.

Table 1: Description of SciLake's stakeholder groups

Stakeholder group	Description & Comments
Research communities at large	Research communities that wish to navigate vast scholarly content, focusing on high-impact and reproducible research objects.
Scientific communities of neuroscience, cancer research, transportation, and energy	Scientific communities participating in the SciLake pilot program. They will bring their domain-specific data, evaluate and showcase the SciLake services demonstrating their value in the context of domain-specific applications.
Builders/operators of thematic or national SKGs (part of EOSC providers)	Research Infrastructures who are building interlinked catalogues and SKGs. National infrastructures who are building their national portals for discovery and monitoring.
Providers of scholarly communication graphs (part of EOSC/global providers)	3rd party service providers, either commercial or public, who provide cross-domain or domain-specific graphs to research organisations, funders and policy makers.
Experts of standards	Standard development organisations and other organisations and individuals involved in the implementation and use of metadata in scientific research.

Stakeholder group	Description & Comments
Computer Science Communities NLP/AI experts	Experts in: • Graph databases, knowledge graphs, big data structures (performance). • Extracting knowledge from scientific text, e.g., citations to different research products, references, affiliations, semantic/domain specific information. • Domain discipline classifications. • Machine translation mechanisms.
Experts of impact assessment	Research communities and infrastructures who shape the frameworks for assessment of Science, Technology & Innovation studies. SMEs that support research evaluation exercises and studies for policy makers.
Research administrators and policy makers	Requiring access to and analysis of scholarly content to identify trends and assess the impact of research for their organisation.
General public	Facilitate knowledge transfer to generate economic and societal benefits.

SciLake will reach out to the aforementioned target groups early in the project to raise awareness and develop the appropriate services and tools based on each group's needs. Throughout the duration of the project, these groups will be requested to provide insights and feedback on the services and tools that will be evaluated and validated.

4 Engagement strategy

The following set of tables provides an overview of the key aspects of our engagement with each of the target stakeholder groups:

Table 2: Key aspects of SciLake's engagement with each stakeholder group.

Research community at large / Scientific communities of neuroscience, cancer research, transportation, and energy	
Objectives of the engagement	- Promote the importance of SKGs for discovering and monitoring research outputs. - Gather feedback on how they navigate and search knowledge graphs, the researcher journey and what types of queries are of interest. Clarify how domain/discipline semantic information is extracted and how it is used, including impact and reproducibility indicators. - Facilitate fit-for-purpose knowledge discovery via efficient navigation through open publications, data, and research software.
Key messages	- Use open infrastructures to organise scholarly content. - Use navigation effectively to prioritise high impact and reproducible research objects in the vast sea of scholarly content. - Develop innovative and customizable services to cover current and anticipated needs of your research community.
Expected Impact	- Enhanced visibility and value of scholarly content. - Easy, transparent, and effective management of scholarly content. - Increased number of researchers to create, manage, access, and process scholarly content. - Increased number of researchers using open infrastructure.
Organisations to approach	- EOSC Users and research communities via OpenAIRE NOADs. - Pilot participants and their networks: HEVS, CERTH, Karolinska Institutet (KI), University of Oslo, (UiO), ICCS.
How to contact them	- Webinars, newsletter items, direct contacts, leverage the OpenAIRE network, EOSC-A communication channels. - Direct contacts through the network of our partners in WP5.

Builders/operators of thematic or national SKGs (part of EOSC providers)	
Objectives of the engagement	- Facilitate scholarly content acquisition, creation and interlinking of SKGs. Expose them to the tools and services built in SciLake. - Promote collaborations and data exchange. - Enhance the building and operation of the SKGs based on impact indicators.

Builders/operators of thematic or national SKGs (part of EOSC providers)	
Key message	- Create fit-for-purpose and interoperable SKGs (cross-domain, cross-region).
Expected Impact	- Expand the use of SKGs among research and national infrastructures. - Easy interlinking of SKGs and scholarly content. - Easy, transparent, and effective management of scholarly content through the SciLake tools and services.
Organisations to approach	- ESFRI and Science Clusters - National initiatives through the NOADs and EOSC (e.g., Research.fi, NWO, NFDI) - RDA WGs, EOSC Task forces - Pilot participants: CERTH, HEVS, ICCS, KI, UiO
How to contact them	- EOSC Symposium, OSFAIR - Community calls and targeted webinars - API forum - Direct contacts via OpenAIRE NOADs and Science Clusters - Workshops

Providers of scholarly communication graphs	
Objectives of the engagement	- Bring consensus on standards for aggregating and sharing scholarly content. - Facilitate SciLake's services integration in 3rd party workflow following a modular approach and providing open APIs. - Promote collaborations and data exchange.
Key messages	- Democratise scholarly content by facilitating the content acquisition and related services.
Expected Impact	- Increased offer of community driven SKGs. - Extended OpenAIRE data space.
Organisations to approach	- For horizontal graphs: Semantic Scholar, Open Alex, Core UK. - For vertical graphs: dblp, ACM Digital Library, IEEExplore, Knowledge graph for CS (Open University). - RDA WG on Open Science Graphs for FAIRness.
How to contact them	- Direct contacts. - RDA, workshops in OS and tech conferences. - Workshops. - API forum.

Computer Science Communities (big data, graph, AI, NLP) / Experts of standards	
Objectives of the engagement	- Demonstrate infrastructural services and tools developed in SciLake. - Develop collaborations for new algorithms on knowledge graphs, training sets and data exchange. - Promote and share software and tools for graph databases, algorithms for quality assessment, deduplication, information extraction, classification algorithms.
Key messages	- Knowledge graphs for improved automated extraction of information from scientific literature. - Open-source AI tooling. - Offer a powerful and unified API to access and query all Open AIRE dataspace contents.
Expected Impact	- Increased performance for creating SKGs. - Enhanced quality of SKGs produced (disambiguation, enrichment). - Improved reproducibility and replicability assessment functionalities, leveraging improved research object identification and multilingual article segmentation and citation context analysis.
Organisations to approach	- Graph database researchers and providers. - European Language Grid (ELG), European Language Data Space. - Other multilinguality experts (e.g., One AI). - SMEs such as: Semantic Scholar, scite.ai, IRIS.ai, Linknovate, OpenCitations, Scholarcy.
How to contact them	- Direct contacts through the network of our partners in WP4 (SIRIS, ARC, DFKI, OPIX, TUE). - Contribution in scientific conferences.

Impact assessment experts (research communities / SMEs)	
Objectives of the engagement	- Use of the open SKGs as the basis for using data and tools for impact assessment.

Impact assessment experts (research communities / SMEs)

	- Elicit requirements for impact related metrics and indicators and for improved/additional field-specific functionalities for knowledge representation, discovery, and extraction. - Test and demonstrate the developed models and techniques for impact assessment.
Key messages	- Use open databases and tools for impact evaluation. - Ability to use off-the-shelf multiple research impact measures. - Use effective impact analysis services and build value added services.
Expected Impact	- Innovative impact metrics and indicators. - Transparency via proliferation of open data and tools in evaluation exercises.
Organisations to approach	- RISIS Research Infrastructure and its members. - OECD, DG-R&I and other funders. - EOSC and COARA. - Institutes and SMEs that support research assessment (apart from partners OPIX and SIRIS, e.g., Fraunhofer, CWTS, Technopolis Group, RAND, ScienceMetrix, PPMI, etc.). - Research Performing Organisations (RPOs) interested in testing new sources/indicators for research assessment (also GraspOS Horizon Europe project).
How to contact them	- Webinars, social media, events. - Contact research performing and research funding organisations, as well as governmental bodies with the aim of accompanying them in understanding the impact of their research, or the research they support, as well as resource allocation, and develop strategies for maximising the impact of their scientific production. - API Forum with users of the SKGs.

Research administrators and policy makers

Objectives of the engagement	- Demonstration and support of the services created and provided by SKGs. - Analysis of scholarly content to assess the impact of research.
Key messages	- Cooperation with other OpenScience initiatives and infrastructures. - Offering customisable knowledge representation, discovery, and extraction services for scholarly content.
Expected Impact	- Effective identification of research objects and of research trends.

	- Effective reproducibility assessment and tracking.
Organisations to approach	- Governmental bodies. - Higher education and research centres. - Research and innovation funding organisations.
How to contact them	- Direct contacts through the network of our partners, events, social media (LinkedIn).

General public	
Objectives of the engagement	- Provide a coherent presentation of the SciLake's vision and progress.
Key messages	- Raise awareness on Open Science.
Expected Impact	- Amplification of valuable and credible research in scientific communication and news.
Organisations to approach	- N/A
How to contact them	- Project website, social media (Twitter, YouTube).

SciLake will develop engagement activities and a continuous user feedback mechanism to ensure its services' usability, impact, and alignment with the evolving users needs. Through its services, support and training, SciLake will engage EOSC users and achieve a significant level of user engagement from the respective research communities regarding the researcher-based production of scholarly content.

4.1 User engagement

SciLake's engagement strategy is built to achieve a co-design of the use of the SciLake services and to ensure the production of research-based scholarly content.

A key element of the engagement strategy is to establish an API forum for SKGs users (computer science communities and impact assessment experts) and developers (builders and operators of thematic and national SKGs part of EOSC providers, and providers of scholarly communication graphs) to agree on seamless API integration. This forum will be built through OpenAIRE to benefit from the pre-existence engaged audience, to expand the impact and the knowledge about the graphs.

The engagement strategy will revolve around three main themes:

- Establish user feedback mechanisms to ensure usability of SciLake services (see Sec. 4.1.1)

- Support and train EOSC users at using SciLake services (see Sec. 4.1.2)

- Researcher-based production of scholarly content (see Sec. 4.1.3)

4.1.1   User feedback mechanisms

The user experience plays an essential role in SciLake's success. By establishing feedback mechanisms, we can develop strategies to engage users and improve their experience while allowing updates to flow more quickly and effectively. Technical activities will follow an agile early/rapid prototyping methodology with short development cycles. This will facilitate regular user feedback cycles, flexible experiments with various ideas, and increased collaboration.

In addition to pilot testing, we will publish online demos on our website, and we will solicit feedback from a large number and range of external stakeholders (see Table 1) through electronic means:

- customised feedback forms for each service under development, to be published at https://scilake.eu/services
- issue tracker list for each service under development, to which users can register using a ticketing mechanism that will be published at https://scilake.eu/services
- send an email directly to the contact person for each service at the address provided on the specific service page at https://scilake.eu/services

We will use the same feedback mechanisms for the presentations and demonstrations delivered at various events, either webinars and workshops organised by SciLake (see Sec. 5.1.6), or external events (see Sec. 5.1.7).

4.1.2   Support and training of EOSC users

Engaging EOSC users with SciLake's services requires training and support.
Three main webinars will be organised to provide training and awareness on the project outcomes, to be delivered in M26, M28, and M34, respectively. The last webinar will also feature a presentation of a training toolkit that includes self-paced modules, a collection of demos, and other materials that can be reused after the end of the project.

SciLake WP6 is part of the HE INFRAEOSC Communication & Engagement WG, which meets monthly to discuss relevant engagement and communication activities from the HE INFRAEOSC projects to identify synergies and join forces on common activities to maximise impact. We will leverage the EOSC-hub communication channels to reach the broad EOSC user community.

4.1.3 Production of scholarly content

Community governance mechanisms will be implemented to motivate and assist researchers in the production and maintenance of scholarly content, like structured SKGs, domain-specific metadata, and unstructured full-text articles. SciLake will build upon crowdsourcing functionalities from OpenAIRE Explore (and similar services) to facilitate the governance of scholarly content by researchers and the involvement of the respective research communities collecting their feedback in a user-friendly and intuitive way.

A core feature of SciLake is its open and powerful APIs which will provide access to both the raw data lake content and advanced content analysis capabilities which will facilitate community governance of scholarly content and relevant services.

5 Communication and dissemination plan

The communication plan will implement in three main phases:

Phase I - awareness of the issues (M1-M18).

This phase will include interviews on specific topics, to raise awareness about the role and the impact of the project's outcomes, Twitter posts, and newsletter articles to introduce and inform the audience about Scilake outcomes and partners.

Phase II - promotion of individual services/components (M19-M34).

This is a later stage of the communication strategy that will involve the implementation and participation in scientific conferences, training and workshops for broader and more direct communication and engagement of stakeholders but also to maximise and consolidate the knowledge and the impact of the project's outcomes.

Moreover, in this phase, various dissemination toolkits (leaflets, brochures,..) will be used and shared to raise awareness of the impact.

Phase III - promotion of the SciLake solution as it is manifested in the OpenAIRE Graph (M24-M36).

To deliver the project's messages and maximise the impact across the stakeholders identified, SciLake will use multiple dissemination channels and dissemination activities to communicate the ongoing process and future project results effectively. The dissemination channels will serve as the communication medium for interaction and knowledge exchange on the progress and continuity of the project's action and results across all stakeholders.

Community governance is becoming a major topic in scholarly communication. SciLake aims to make publishing workflows and infrastructures more accountable to a wider range of stakeholders by working on:

- Open governance to services and the OpenAIRE Graph.
- Organisational structure & workflows.
- Mechanisms through which organisations (partners and 3rd parties) can participate in the governance of services, and under certain conditions, provide services.

5.1 Communication and Dissemination channels

Communication channels will enhance the awareness of the key target groups through continuous updates about the activities and action of the SciLake project.

5.1.1 Project Website

The https://SciLake.eu/ website is the primary communication and dissemination channel for the SciLake project. The website will serve as the central channel to include all project updates, actions, results, and outcomes. Sharing all activities, services, and results via the website will ensure that SciLake users will be informed continuously, transparently, and efficiently.

At the time of writing, visitors and users can find some general descriptions of the project, the objective, and the methodology of the SciLake project.

The website will provide references to the following project outcomes:

- All public deliverables
- All case studies included in SciLake
- The services which will be provided by the project
- News and the following events that will take place

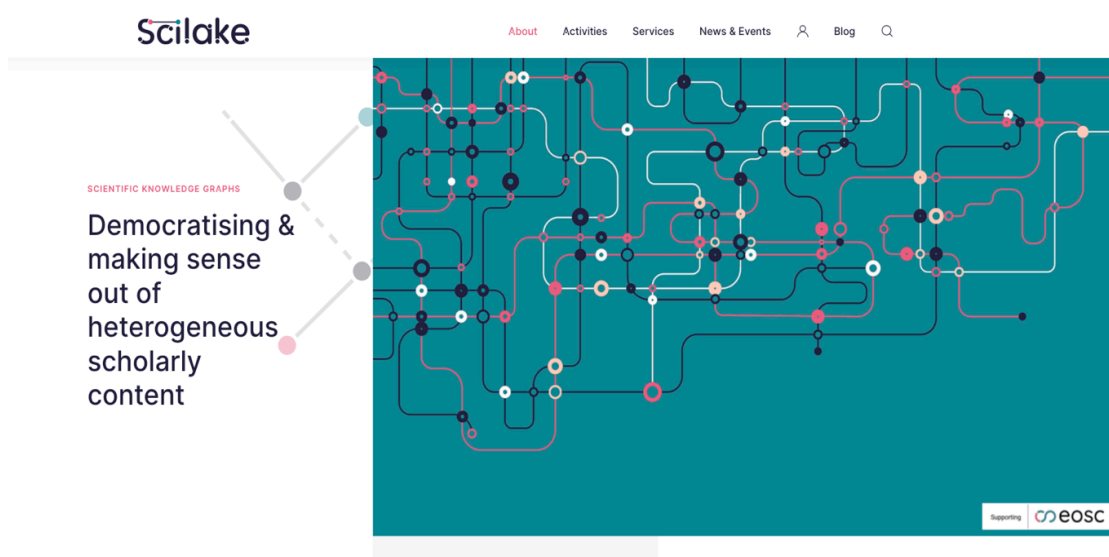*Figure 1: An image snapshot of SciLake's website homepage.*

Figure 1 shows an up-to-date image of SciLake's website homepage. More material and updates of the website will be offered throughout the project as the website will be promoted constantly via the rest of the dissemination channels. Figure 2 shows the primary structure of the website.
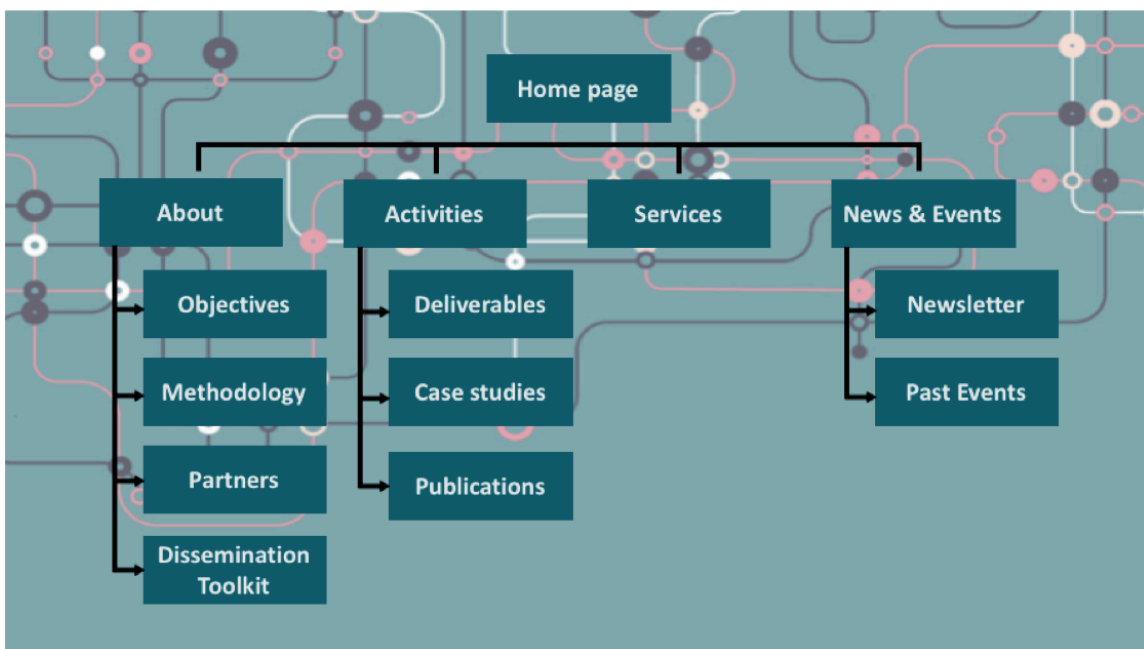

Figure 2: Structure of the website menu.

5.1.2   Social Media

The SciLake project will be active on Twitter, LinkedIn, and YouTube, providing various gateways for the audience to get acquainted and interact with the partners, content, and updates. The value of social media as channels for communication and dissemination tools will be crucial throughout the project to reach out to a broader audience and raise awareness about the SciLake's progress and outcomes.

In order to achieve the best outcome and create a stable and engaged audience, SciLake will rely on partners' networks to spread all outgoing messages and to maximise the project's impact. Posts and updates on social media will be constant and continuous. The aim of the social media channel will be:

Twitter: Maximise the project's impact and visibility to a wider audience and raise awareness about its actions and activities. This channel is the most frequently updated, as it is ideal for creating short posts for project updates. Project partners will be tagged and forward the posts. Some relevant hashtags that will be included and used are #OpenScience, #ScienceKnowledgeGraph, #SKG, #OpenAIREGraph, #EOSC, #ScientificLake, #SciLake, and so on. Moreover, short video material will be included, such as interviews with the partners, promotional videos and event recordings. The first project's months aims at least one weekly/bi-weeklyy update on the current actions. The activity of this channel will increase exponentially once the project outcomes are finalised and can be published.

LinkedIn (OpenAIRE): Refer to a more professional network aiming at reaching out to policy makers. LinkedIn is a more professional channel referring to and expanding visibility among related communities. SciLake project will be promoted by the OpenAIRE account to automatically refer and aware a wide audience maximising from the beginning the spread and awareness of the impact of the project. Through the use of the community of the OpenAIRE account, SciLake will share with an already existing, broad audience 2 to 3 times per month. As the project results become finalised, the frequency may increase.

YouTube (OpenAIRE):

Starting from M20, through the Youtube channel of OpenAIRE, the project will share at least five videos presenting various topics like the followings:

- Interviews from the pilots and the partners
- The outcomes of the project
- Use of SKGs by stakeholders
- Tutorials for training purposes

5.1.3   Newsletter

The project newsletter will be published featuring the most important updates and results, in a summarised way, for a wider community complementing more technical information to a

more specific and targeted audience. For each version of the newsletter, a contribution from all partners will be required to ensure that all crucial activities will be included in the content. In year 1 the newsletter will be published in M7 and M12. The frequency of the newsletter will be re-evaluated in year 2 when more information will be available, and there might be the need to share information via this channel more frequently.

The newsletter will be hosted in a dedicated section of the project website (https://scilake.eu/news), where an email subscription option will also be available. To engage subscribers and maximise the outreach of the newsletter, the SciLake communication team (WP6) will systematically rely on the existing networks of partners of the consortium. The project highlights will also be featured in partners' newsletters (OpenAIRE, European Language Technology, EBRAINS).

5.1.4   Success stories

Every pilot and partner of this project will describe a success story promoting their results. This promotion will encourage and engage new stakeholders to exploit the outcomes of Scilake further and maximise its impact and dissemination. These stories will be published on the project website and through social media (Twitter, LinkedIn) posts.

5.1.5   Publishing project results

The project objective, strategies, and plan to achieve and maximise the impact include publishing the project results on software and open data repositories. Adopting the Open Science practices, all SciLake's software components will be released as open-source code on software repositories (e.g., GitHub), and all datasets produced will be available via open data repositories such as Zenodo. All public deliverable reports will be published on the SciLake Zenodo community and will be referenced in the project website.

Table 3 lists the publications so far.

The expected number of publications in the project lifetime is 20.

Table 3: List of publications

Title	Where	Authors
Piloting topic-aware research impact assessment features in BIP! Services	Extended Semantic Web Conference (ESWC), 2023	S. Chatzopoulos, K. Vichos, I. Kanellos, T. Vergoulis

Atrapos: Real-time Evaluation of Metapath Query Workloads	In Proceedings of the ACM Web Conference (WWW), 2023, doi.org/10.1145/3543507.3583322	S. Chatzopoulos, T. Vergoulis, D. Skoutas, T. Dalamagas, C. Tryfonopoulos, P. Karras

5.1.6 Partner's communication and dissemination channels

The project partners' communication channels and circuits can be exploited to reach a wide audience. The SciLake communication and dissemination team will provide continuous support and materials to promote the project and ensure its visibility through the partners' communication channels.

5.1.7 Events organised by SciLake

The project will organise events to share its outcomes with diverse partner groups and reach various stakeholders. The events will either be internal meetings of all SciLake partners or workshops and webinars intended to foster and share knowledge and solutions with key stakeholder groups. These events will all provide essential input for the successful completion and implementation of the project.

Table 4 lists the events planned with these objectives. For each event, more information will be provided at the appropriate time.

Table 4: List of SciLake events.

Event	Type	Aim	Target Group(s)	Timeline	Place
Kick-off meeting	Internal SciLake meeting	Introduction of members and groundwork preparation	All SciLake partners	23-24 January 2023	Athens, Greece

Plenary meeting	Internal SciLake meeting	Review the progress made, discuss challenges and plan next steps	All SciLake partners	9-10 November 2023	Barcelona, Spain & Online
Workshop 1: Presentation of the first integrated system	Workshop	Gather feedback on the system	SKG builders and providers, computer science community	Delivered in M19	Format & Location to be confirmed
Workshop 2: Presentation of the pilot setup and evaluation of the methodology	Workshop	Showcase SciLake services and gather feedback in the context of domain-specific applications	Scientific communities in neuroscience, cancer transportation, energy, and research communities at large	Delivered in M24	Format & Location to be confirmed
Webinar: Training	Webinar	Providing training & awareness on project outcomes	Members of all SciLake target groups	Delivered in M26	Online
Webinar: Training	Webinar	Providing training & awareness on project outcomes	Members of all SciLake target groups	Delivered in M34	Online
Workshop 3: Finalisation of the services	Workshop	Final validation of the services	Members of all SciLake target groups	Delivered in M34	Format & Location to be confirmed
Webinar: Training	Webinar	Presenting the training toolkit	Members of all SciLake target groups	Delivered in M34	Online

Final SciLake Plenary meeting	Conference	The final event of SciLake aimed at presenting all key projects outputs	All SciLake partners and representatives of SciLake target groups	Delivered in M35	In–person: Location to be confirmed

5.1.8 External Events

To ensure the successful promotion of project results and raise awareness about SciLake's main aims and outputs, project partners will be encouraged to attend key events organised by external parties/organisations. WP6 will encourage project partners to attend and assist partners in participating in those events. A list of relevant upcoming events has been created listing important deadlines (see below the Monitoring section of the plan), call for proposals' details, major themes, and key stakeholder groups targeted at each event. By participating in external events, the relevant scientific communities can get acquainted with the project, monitor its progress and join future SciLake actions. A list of possible events could include (but is not limited to) the following: conferences, workshops, OS focused events, exhibitions, Horizon Europe, and EC events. Table 5 provides an indicative list of relevant major events to be followed and attended by SciLake partners (as the project progresses, more will be included).

Table 5: Preliminary list of events attended or to be attended by SciLake partners.

Event	Type	Organiser	Target Group(s)	Timeline	Place
The Web Conference 2023	Conference	TheWebConf	Computer science experts, research community at large	30 April - 4 May 2023	Austin, Texas
Extended Semantic Web Conference (ESWC), 2023	Conference	ESWC	Knowledge graph experts, computer science communities	28 May - 1 June 2023	Herson-issos, Greece

ITAEOSC2023	National event	CNR	Research community at large, Open Science experts	June 5, 2023	Rome, Italy
OS FAIR 2023	Conference	OpenAIRE	Research support staff, OS experts, RI management National eInfras, R&I policy makers & policy officers	September 25-27, 2024	Madrid, Spain
European Conference on Computational Biology (ECCB)	Conference	ISBM/ECCB	Research communities of computer science, molecular biology, mathematics, and statistics	July 23-27, 2023	Lyon, France & online
European Climate and Energy Modelling Platform (ECEMP)	Conference	ECEMF	Research community at large Scientific communities of energy research	October 04-06, 2023	Virtual
14th European Breast Cancer Conference (EBCC)	Conference	EBCC Council	Research community at large Scientific communities of cancer research	March 20-22, 2024	Milan, Italy

5.1.9   Training plan

Training activities will be developed during this project to show stakeholders how to build and use interlinked SKGs and smart impact discovery and reproducibility assistant services.

Technical providers will prepare training material (documentation, demos) to be presented in one webinar in M24, and two webinars in M34, accompanied by tutorial videos recorded on Youtube using OpenAIRE's channel. All the training material will be structured into an online training course (Moodle) hosted on the OpenAIRE learning platform, OpenPlato (www.openplato.eu), and presented in the final webinar in M34. The course will also be referenced in the EOSC training platform.

Table 6 outlines the topic, objective, format, and training target.

Table 6: Training plan.

Topic	Objective	Format	Target groups
Knowledge Graph creation assistant	Learn how to create a collection of SKGs for heterogeneous scholarly content	Webinar section (M24) Video tutorial (M24) Module in self-paced course (M34)	Builders/operators of SKGs
Data interlinking	Learn how to use tools and methods for data interlinking in SciLake's SKGs	Webinar section (M24) Video tutorial (M24) Module in self-paced course (M34)	
Data Lake search and navigation	Learn how to search and browse through the contents of the Scientific Lake	Webinar section (M24) Video tutorial (M24) Module in self-paced course (M34)	Research community at large Graph providers Computer science communities
Smart impact discovery service	Understand AI-assisted tools for effective identification of research objects, research impact assessment and trend identification	Webinar section (M34) Video tutorial (M34) Module in self-paced course (M34)	Research community at large Impact assessment experts Research administrators/policy makers Computer science communities

	Understand AI-assisted tools for automated assessment of research reproducibility	Webinar section (M34) Video tutorial (M34) Module in self-paced course (M34)	
Smart reproducibility assistant service			
Open API	Learn how to use the SciLake open API	Webinar section (M34) Video tutorial (M34) Module in self-paced course (M34)	Builders/operators of SKGs Graph providers Research administrators/ policy makers Computer science communities

Pilot participants will receive additional training on SKG use for their specific case studies which will be planned in Deliverable D5.1 (M23).

5.1.9.1 TRAINING ASSESSMENT

Training assessment ensures that participants receive optimal training and enhance their knowledge and skills. The assessments will help identify areas where stakeholders may need additional support, as well as evaluate the effectiveness of the training.

In practice, participants in each activity will be asked to provide feedback and these data will be collected for further evaluation to optimise the process. A feedback form will be shared through the registration emails for webinar participants. A link to the appropriate contact information will be displayed for YouTube video tutorials to allow users to send free feedback or requests. The online training course on OpenPlato will include a survey about the training received that participants must complete upon completion. Participants will also have the opportunity to post questions and comments at any time in a dedicated forum on OpenPlato.

5.1.9.2 SUPPORT MATERIAL

Table 7 provides an overview of the main supporting material that will be used for training. This material will be regularly updated for optimization purposes and whenever new services and tools emerge.

Table 7: Training supporting material.

Guides	Use cases	Video tutorials
On building / using interlinked SKGs How to search and navigate through the Scientific Lake On using the open self-documented API On using trend identification analysis	Of using SKGs by the different stakeholders and pilots Of application of multi-perspective of scientific impact Of application of trend identification analysis	Produce videos for training and demonstration purposes
Factsheets	**FAQs**	**Webinar and Community Calls**
References on the impact measures References on Scilake services and tools	Answers and comments to the provided services and tools	Presentation of the participants regarding domain general and specific SKGs, impact measures to valuable research and recording of the calls

5.2 Communication and Dissemination tools & materials

5.2.1　Dissemination toolkit

For the consistent presentation of the SciLake brand, a dissemination toolkit has been created and published on the SciLake website.

- Project logo

Figure 3: SciLake logo.

- A PPTslide deck template to be used consistently in all partner presentations.

- Two virtual backgrounds (light/dark themes) for online meetings/webinars/workshops.
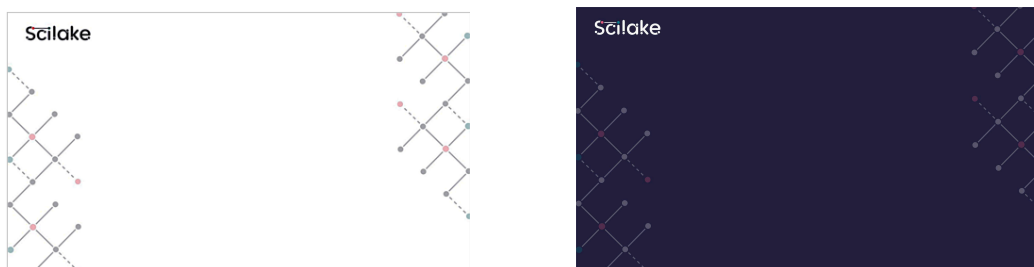
Figure 4: Virtual backgrounds for online event participation.

- A SciLake poster introducing the project.

Figure 5. SciLake introductory poster.

5.2.2   Project brochure

To introduce the objectives and the aims of the project in a more marketable way, a brochure will be produced. Through this brochure, all essential information and services provided will be presented to promote all the key messages of the project. The brochures will be printed and used as promotional material during events.

5.2.3   Standard project presentation

To maximise the impact of SciLake project, a standard project presentation that describes in a short and simplified way all the information and the objectives of the project and introduce the partners will be shared. This presentation can be modified to fit specific audiences and tailored to each occasion.

5.2.4   Project templates

To maintain a uniform way of presenting and reporting the project's outcomes, several templates will and have been created so they can be used by all partners from all WP and present the results in a consistent manner.

The templates include (but are not limited to):

- A PowerPoint presentation (short description of the project and the case studies (see Annex 1)
- Standard word document for deliverables (see Annex 2)
- Standard word document for WP meetings

5.2.5   Interviews

To maximise the impact and be informed about the activities taking place in all WPs, interviews with technology providers and pilot program representatives will be conducted using a Q&A format. The interviews will aim to gain a deeper understanding of the role of each partner and the outcomes of SciLake. A similar communication strategy will be followed throughout the project with external stakeholders. Doing so will increase awareness and build a relationship based on continuous information/feedback exchange to optimise the results.

6 Exploitation

This document presents a preliminary exploitation plan that will be revised once the final project's outcomes are determined. All SciLake partners will participate in an internal workshop between M14 and M16 to discuss and finalise the project's expected results and impact. On this occasion, we will review the preliminary exploitation plan (Sec. 6.1) to enhance SciLake's added value in terms of open research information, as well as plan how to connect with existing and ongoing initiatives (from COARA to OpenAlex). A report of the exploitation activities will be presented in Deliverable D6.2 (M36).

6.1 Preliminary exploitation plan

Our exploitation strategy will be driven largely by the initiatives and plans pursued by our partners OpenAIRE and SIRIS Academic in the areas of open science and policy-making, respectively.

OpenAIRE is actively cooperating with other open infrastructures, such as Crossref, OpenCitations, and Open Abstracts, to develop solutions for fostering and providing open access metadata about scholarly outputs. OpenScholar is one of the activities of this group that aims to identify the challenges associated with replacing the Microsoft Academic Graph (MAG) with a similar but community-governed service, so as to ensure long-term sustainability is in the hands of scientists and other related organisations. To accomplish this goal, SciLake aims to provide the tools and office procedures that are required to operate a service that is able to meet many (ideally all) of such requirements, both technologically and in terms of governance.

SIRIS Academic is a research-driven consulting company specialised in science and innovation policy-making. SIRIS regularly conducts research portfolio analyses and data-driven impact assessments. SIRIS has established a number of partnerships with research funders and researcher associations, particularly in the fields of neuroscience and cancer research. Within this context, the tools developed by SciLake will greatly enhance the services provided by SIRIS: specifically, SIRIS' consulting arm will continually utilise the results of the "Smart Impact Analysis Service" and the "Smart Reproducibility Analysis Service" to gain a more detailed understanding of scientific impact within specific research areas. In parallel, by developing some of the technologies to be applied in the Smart Impact and Smart Reproducibility analysis services, SIRIS will be able to refine its quantitative methodologies and machine learning tools for analysing textual documents related to Science, Technology and Innovation. In addition, participation in the pilots and the co-design of their respective services will provide SIRIS with a framework for formalising participatory protocols in order to define semantic perimeters for specific research areas.

The developed services will ultimately be integrated into the **EOSC core service platform** according to predefined access policies, a tiering model, and pricing that will be determined to guarantee their long-term sustainability.

Aside from OpenAIRE and SIRIS's major actions, all SciLake partners will contribute to the exploitation of the project's results. The following tables present the role and initial expected outcomes of the project for the partners involved in technical development and science case studies.

Athena Research Center (ARC)

Role:
Project coordinator. Actively participate in the design of the SciLake architecture. Lead the design and implementation of the smart impact-driven discovery service. Participate in the design and implementation of the smart reproducibility assistance service. Work together with the pilots for the requirements acquisition and analysis.

Expected outcomes:
Deliver components that contribute towards the implementation of an EOSC-integrated and customisable smart impact-discovery service that uses advanced technologies to estimate research impact of research artefacts according to various perspectives producing a set of indicators that can facilitate the identification of research trends and valuable research objects. This work will build upon the BIP! Services (https://bip.imsi.athenarc.gr/site/home). Deliver components that contribute towards the implementation of an EOSC-integrated and customisable smart reproducibility assistance service. Notable examples are technologies related to topic/field extraction, automatic identification of research artefacts, link recommendations among research artefacts, text analysis of multilingual content.

Centro Nazionale delle Ricerche (CNR)

Role:
Initiate and drive SciLake activities through a phase of collection of the requirements followed by a phase of co-design of the services to be developed, done together with the pilot partners. Design SciLake's architecture and integrate all the contributions provided by work packages WP2, WP3, WP4, taking care of the FAIR principles when dealing with data acquisition and data generation.

Expected outcomes:
Deliver tools to facilitate content acquisition, and implementation of FAIR protocols for SKGs data catalogues.
Develop an open and self-documented API to provide access to the SKG workflow, as well as access to the raw content of the Scientific Lake.

Technische Universiteit Eindhoven (TUE)

Role:
Lead the design and implementation of the data lake service for hosting, managing, and querying heterogeneous (e.g., different scientific fields, formats, languages) scholarly content.

Expected outcomes:
Develop automated and semi-automated tools to facilitate the creation and data interlinking in SKGs. Make effective use of the research work of the Eindhoven DB group and obtain novel solutions for challenges encountered in use-cases of the project.

Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI)

Role:
Develop Natural Language Processing (NLP) and Language Technology tools for scholarly information processing, in particular for the identification and extraction of specific information from scientific articles, as well as for the overall detection of research paper structure. These technologies will be made available via the cloud platform European Language Grid (ELG).

Expected outcomes:
Develop a fully working scholarly information processing prototype that is orchestrated and deployed through the European Language Grid and that is able to provide meaningful services to the SciLake project as a whole. Additionally, DFKI would like to establish a bridge between SciLake – as an EOSC project – and the project NFDI4DataScience and AI – which is one of the projects that develop the National German Research Data Infrastructure (NFDI).

OPIX P.C. (OPIX)

Role:
Contribute to the design and implementation of the smart impact-driven discovery service, the smart reproducibility assistance service and the development of an appropriate suite of APIs facilitating end-users in exploiting SciLake services.

Expected outcomes:
Build SciLake services and provide seamless access through an integrated API to EOSC service providers and users and the research community (including research policy makers) at large.

HEVS

Role: Lead the <u>Energy</u> research community pilot. Provide energy research database and SKG validation.
Expected outcomes: Improve the Enermap service through SciLake, so that users can easily access data, explore the subject of energy, identify methodologies and protocols, and access relevant scientific literature.

Universitetet i Oslo (UiO)
Role: Provide a <u>Neuroscience</u> database and expertise of the EBRAINS Research Infrastructure and the Knowledge Graph, to the SciLake infrastructure.
Expected outcomes: Connect the EBRAINS knowledge graph to other FAIR initiatives to find robust ways to measure data impact, and to understand the relevant parameters for making data available for meta-analyses.

Institute of Communication and Computer Systems (ICCS) Centre for Research & Technology, Hellas (HIT/CERTH)
Role: Create a domain-specific knowledge graph using related internal or external data sources by participating in <u>Transport</u> Research's pilot project.
Expected outcomes: Provide AI-assisted services to the transport research sector to assess the significance of relevant research (whether it is impactful or reproducible), to discover associations among technical specifications, published standards, draft standards and their appearance in research work.

○

Karolinska Institutet (KI) Centre for Research & Technology, Hellas (CERTH)

Role:
Lead the <u>Cancer</u> pilot study, with the purpose of exploring the potential of using knowledge graphs for cancer research in general, and for Chronic Lymphocytic Leukaemia (CLL) in particular. The pilot will evaluate and test new elements and services, such as compiling evidence for a biomarker based on user-defined data and literature, generating the highest-ranking path that connects research objects to stratify patient cohorts.

Expected outcomes:
- Use of the Scientific Lake service to directly interface EOSC resources to the relevant databases and construct a domain-specific KG, representing the domain which relates to the respective research questions (incl. patient stratification based on disease prediction).
 - Use of the impact analysis service and reproducibility analysis service to prioritise reading in the context of the clinical interpretation of genomic variants that have been identified as interesting via NGS analysis.

7 Monitoring

Regularly monitoring the impact and success of all communication, engagement and dissemination activities is crucial in evaluating the project's actions and thus ensure the delivery of the results while keeping all key target groups engaged. As the communications, engagement and dissemination plan is envisaged to be a living document that will be regularly updated during the project's lifecycle, having a monitoring system in place will help towards updating, revisiting, and optimising the approaches adopted by SciLake. Monitoring of the impact of these activities will help determine the need for further or complementary actions to achieve the promotion, visibility and uptake of project results.

Three main lists have been created to monitor and provide updates on the project's activities. These files have been uploaded on the collaborative shared space of the project and will be regularly updated both by WP6 and all partners.

SciLake Events

Event	Date, Place	Purpose	Number of attendees	Relation to WPs and Tasks	Supporting documentation (links)	Targeted stakeholders	Additional information

Scilake

External events to be attended by SciLake

Event	Date, Place	Submitted?	Participation (workshop, presentation, poster, etc)	Deadlines, details for the call for proposals	Relation to WPs and Tasks	Supporting documentation (links)	Targeted stakeholders + estimated numbers	Additional information

Scilake

Relevant projects, possible collaborations, synergies with SciLake

Project	Relevance to Scilake	Possible collaboration	Relation to WPs and Tasks	Supporting documentation (links)	Targeted stakeholders + estimated numbers	Additional information

Figure 6: Event monitor spreadsheets.

8 Conclusions

It is important to ensure that the project achieves its goals with a positive impact on research. To become successful, it is imperative that the project understands the needs of researchers and reaches them through the right communication channels. The aim is to take the appropriate actions to engage researchers throughout the design and implementation phases. To achieve this goal, the Engagement and Communication Strategy and Plan provides a coherent framework and guidance to structure relevant activities and presents the planned engagement and communication activities.

This deliverable will be the backbone for all work packages' engagement and communication activities throughout the SciLake project, with WP6 monitoring that alignment. WP6 has therefore assigned at least one of its members to attend the meetings of each work package. Also will raise awareness of SciLake's plans and ambitions in the community by synchronising activities and agreeing on the engagement plans.
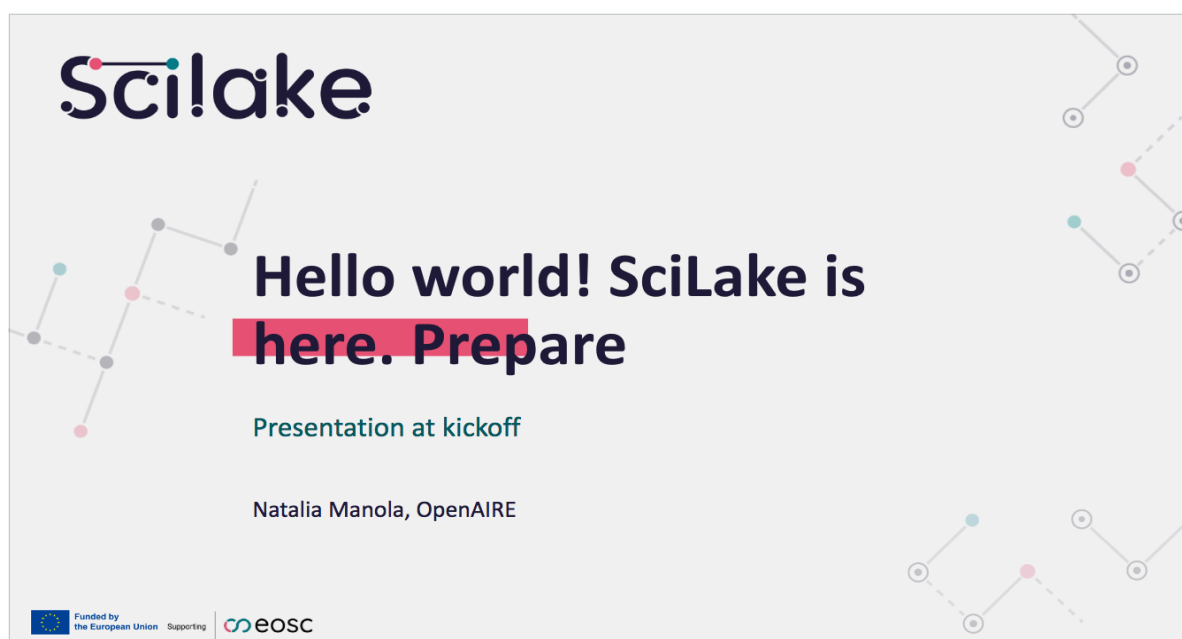
9 References

10 Annexes

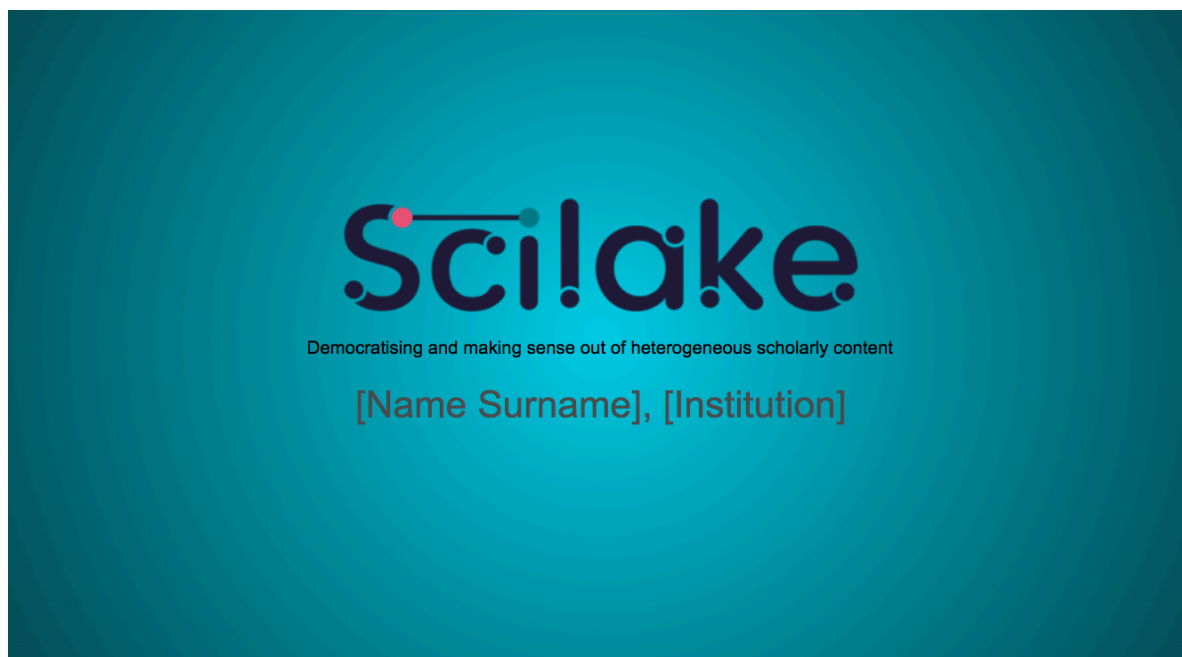## 10.1 Annex 1

Figure 7: Short presentation of SciLake.

Figure 8: Short presentation of SciLake case studies.

10.2 Annex 2

Scientific Lake

Deliverable xx

Due Date of Deliverable	
Actual Submission Date	
Work Package	
Tasks	
Type	
Approval Status	
Version	
Number of Pages	

Abstract

Figure 9: Template for deliverables' preparation.