Communication and Dissemination Plan

Submission Date	31.03.2024
Due Date	31.03.2024
Version	0.1
Deliverable Title	Communication and Dissemination Plan
Deliverable No.	2.1
Lead beneficiary	IBL-PAN (BUT)
Contributing WP	2
Funding id	

Project Full Title	Making Software FAIR: A machine-assisted workflow for the research software lifecycle
Project Acronym	SoFAIR
Project No.	
Start Date	2024
End Date	2025
Duration	24
Project Website	
Authors	Patryk Hubar-Kołodziejczyk, Tomasz Umerle, Nikodem Wołczuk
Abstract	This task focuses on designing the SoFAIR communication, dissemination plan, and exploitation plan, where the strategy for the project's communication and dissemination is outlined with a clear division of responsibilities between partners. The communication and dissemination plan will identify the target stakeholders, in particular research performing and funding organisations (RFPOs), repositories, libraries, research managers, working groups managing relevant standards (OpenAIRE, RIOXX, NISO, RDA, CLARIN, DARIAH, etc.), scholarly infrastructures (including publishers). Through channels of communication – such as the project website, social media accounts and participation and organisation of events – the project's output will be communicated and disseminated in a manner that will facilitate its uptake by the community of varied stakeholders, including different service providers and researchers.

Version and Revision History

Version	Date	Author/Reviewer/Contributors	Comments
0.1	29.02.2024	Patryk Hubar-Kołodziejczyk, Tomasz Umerle, Nikodem Wołczuk	First draft
0.2	14.03.2024	Matteo Cancellieri, Alain Monteil	Reviews
1.0	28.03.2024	Patryk Hubar-Kołodziejczyk, Tomasz Umerle, Nikodem Wołczuk	Final version

List of Acronyms

C&D	Communication and Dissemination
CDP	Communication and Dissemination Plan
D	Deliverable
FAIR	Findable, Accessible, Interoperable, and Reusable
PIDs	Persistent identifiers
ML	Machine learning
NGOs	Non-Governmental Organisations
KPIs	Key Performance Indicators

Table of Contents

Executive Summary

The SoFAIR Communication and Dissemination Plan (CDP) is a foundational document for the Making Software FAIR project, a pioneering initiative to improve the discoverability, attribution and reusability of open research software through the FAIR (Findable, Accessible, Interoperable, and Reusable) principles. This plan carefully outlines a strategic approach to communicating and disseminating the results of the project, with a focus on increasing the visibility and adoption of its outputs by a diverse community of stakeholders, including research performing and funding organisations, repositories, libraries, research managers and scholarly infrastructures. **CDP** emphasises a multi-channel approach, using the project website, social media, participation in and organisation of events, and publication of scientific results to engage with and inform the project's target audiences. It specifies the project's commitment to internal and external communication strategies designed to promote the project's innovations, facilitate stakeholder engagement and ensure the wide dissemination and uptake of its results. **Key stakeholders identified** include libraries, open repositories, research infrastructures, publishers, research institutions, individual researchers, NGOs, funders, the general public and policy makers. The CDP details a tailored **engagement strategy** for each of these groups to ensure the delivery of relevant and impactful messages adapted to their specific interests and needs. The document also outlines a series of **Key Performance Indicators (KPIs)** to measure the effectiveness of the communication efforts, including the number of blog posts, publications, webinars and social media interactions planned and executed throughout the life of the project.

1 Introduction

1.1 Project Description

A key issue hindering the discoverability, attribution and reusability of open research software is that its existence often remains hidden within the manuscript of research papers. For these resources to become first-class bibliographic records, they first need to be identified and subsequently registered with persistent identifiers (PIDs) to be made FAIR (Findable, Accessible, Interoperable and Reusable). To this day, much open research software fails to meet FAIR principles and software resources are mostly not explicitly linked from the manuscripts that introduced them or used them.

This project will extend the capabilities of critical and widely used open scholarly infrastructures (CORE, Software Heritage, HAL) and tools (GROBID) operated by the consortium partners, delivering and deploying an effective solution for the management of the research software lifecycle, including 1) ML-assisted identification of research software assets from within the manuscripts of scholarly papers, 2) validation of the identified assets by authors, 3) registration of software assets with PIDs and their archival.

The solution will be optimised for deployment over open content available through the global network of open repositories aggregated by CORE (core.ac.uk), which constitutes over 32 million full texts and 250m+ metadata records from over 10k repositories currently the world's largest collection of open access documents. Our ML software for extraction and disambiguation of software assets will be realised as an extension of the state-of-the-art GROBID tool. We will build on established protocols, such as OpenAIRE Guidelines v4.0, RIOXX v3 and Codemeta, to encode information about software assets and their links to research manuscripts establishing an interoperable and extensible workflow connecting open repositories (represented by HAL), aggregators (represented by CORE) and software archives (represented by Software Heritage). The efficacy of the developed tools and workflow will be validated in three use cases: 1) a life sciences demonstrator (for Europe PMC), 2) a multi-disciplinary demonstrator for institutional repositories (represented by HAL) and a 3) digital humanities case study (with links to DARIAH and EOSC).

SoFAIR will improve and semi-automate the process for identifying, describing, registering and archiving research software, ensuring it has received a Software Heritage persistent identifier (SWH-ID). The solution extensively builds on enhancing existing widely used open source tools (GROBID) and open scholarly infrastructures (CORE, Europe PMC, HAL, Software Heritage), which are operated by the consortium members. This ensures fast and wide adoption of the project’s outputs across the global scholarly ecosystem of open repositories, offering tangible pathways to impact.

The project closely aligns with the FAIR agenda, Horizon Europe initiatives on open research data and software, and European infrastructures including EOSC, CLARIN and DARIAH. It will apply a number of open protocols and recommended scholarly metadata standards such as, but not limited to, CodeMeta, OpenAire Guidelines, Rioxx and Signposting, to describe identified, enriched and curated research software assets as interoperable metadata and connect them to research manuscripts through explicit links.

The project will also create a new reference manually annotated gold dataset for the training of machine learning tools for automatic identification, enrichment and disambiguation of software mentions and advance the state-of-the-art on this task.

By lowering the barriers for research software to be made FAIR, this project will drive a step-change in the way research software can be found, accessed and referenced. This will facilitate not only software reuse and reproducibility but will also contribute to incentivising the creation of research software.

1.2 Deliverable Description

A Communication and Dissemination Plan is a strategic document that is an integral part of the management and implementation of projects, particularly those with research, innovation, or community engagement components. It outlines the methods and actions that will be used to effectively share the project's results, knowledge and outputs with a wider audience, including stakeholders, potential users and the general public.

The main objective of this plan is to maximise the impact and visibility of the project's results. It achieves this by identifying key messages, target audiences and the most appropriate channels and tools for dissemination and communication. The plan is tailored to ensure that the information reaches the relevant stakeholders in an accessible and engaging way, facilitating the wide uptake and exploitation of the project results.

1.3 Preparation & Methodology

The following steps were taken in the preparation of the **Communication and Dissemination plan**:

- **Objectives definition**: Establishing clear, SMART objectives for the communication efforts, ensuring they were directly aligned with the overarching goals of the project,
- **Target audiences identification**: A comprehensive analysis was conducted to understand the characteristics, needs, and preferences of each target audience, allowing for the creation of tailored communication strategies,

- **Communication channels and tools selection**: Based on the preferences of the target audiences and the nature of the content, the most appropriate channels and tools for communication were chosen to ensure a wide and effective reach,
- **Activities and timeline plan**: Specific communication and dissemination activities were outlined, along with their respective timelines, creating a coordinated roadmap for the execution of the communication strategy,
- **Responsibilities assignment**: Clear roles and responsibilities were defined within the project team, facilitating the effective management and implementation of the communication plan,
- **Monitoring and evaluation activities:** The plan incorporated mechanisms for assessing the effectiveness of the communication efforts, with performance indicators and metrics set for ongoing evaluation and necessary adjustments.

2 Stakeholder Analysis

The communication and dissemination activities of the SoFAIR project will be tailored to meet the different needs and preferences of a wide range of key stakeholders. These include cultural heritage entities such as open repositories and libraries, which form the backbone of scholarly resources. Research infrastructures and publishers play a key role in the dissemination and accessibility of scholarly work, while research institutions, universities and individual researchers, including working groups, drive the **academic** pursuit of knowledge. On the other hand, **non-academic stakeholders** include public authorities and funding agencies, which are crucial for policy and financial support. Policymakers and various organisations and individuals outside the traditional academic sphere are also an important focus of the SoFAIR project's engagement strategy. These stakeholders play a key role in translating academic research into practical societal benefits and in shaping the environment in which research and innovation take place. By recognising the different roles and interests of these academic and non-academic stakeholders, SoFAIR aims to make the best use of existing communication channels and networks. This approach not only facilitates the effective dissemination of messages but also underlines the project's commitment to fostering a collaborative ecosystem. Such an ecosystem promotes the sharing of knowledge, resources and best practices across the divide, ensuring that the project's results are widely recognised and integrated into both academic research and wider societal application

3 Action Plan

3.1 Strategic Approach

The main objectives for engaging SoFAIR stakeholders is to **inform** about the project's goals and results, **share and consult** methods and approaches and **raise awareness** around the importance of recognising software as a key scientific resource.

The engagement of SoFAIR stakeholders revolves around three critical areas of communication to enhance the project's impact and collaboration:

- **Information about the consortium**:
    - **Objective**: To present the partners and individuals involved in the project, highlighting their roles, expertise, and the potential achievements through interdisciplinary collaboration.
    - **Importance**: This aims to not only showcase the diverse capabilities and contributions of each consortium member but also to emphasise the value and impact of interdisciplinary cooperation in achieving the project's objectives.
- **Information about methods**:
    - **Objective**: To share the methodologies, frameworks, and tools developed or employed by the project, encouraging adaptation by the wider community.
    - **Importance**: Promotes the adoption of best practices in software development and maintenance within scientific research, contributing to the project's legacy and the broader research field's advancement. This also fosters the promotion of software citation methods and the adoption of standards for ensuring high-quality software metadata, thereby enhancing the reproducibility and traceability of research outputs.
- **Information about results**:
    - **Objective**: To disseminate the project's findings, outcomes, and broader impacts, highlighting the significance of software as a scientific resource.
    - **Importance**: Aim to raise awareness of the critical role of software in scientific discovery and advocate for its recognition, sustainability, and proper acknowledgement within the research ecosystem.

Tab. 1. Stakeholders description and communication areas

Stakeholder	Description	Message – points of emphasis
Libraries and open repositories	Entities managing collections of digital and print resources.	Inform about results
Research infrastructures *CLARIN, DARIAH*	Networks provide digital tools and resources for research.	Inform about consortium Inform about methods Inform about results
Publishers	Organisations responsible for the dissemination of research findings.	Inform about results
Research institutions	Academic or non-academic bodies conducting research activities.	Inform about methods Inform about results
Individual researchers and working groups	Individuals or groups focused on specific research areas or projects.	Inform about methods Inform about results
NGOs	Non-governmental organisations with interests in specific research impacts.	Inform about results
Funders	Organisations or individuals providing financial support for research.	Inform about consortium Inform about results
General public	The wider community with a potential interest in the project's outcomes.	Inform about results
Policymakers	Individuals or bodies involved in the creation and enforcement of policies.	Inform about consortium Inform about results

Stakeholder engagement in the SoFAIR project is planned to be multi-faceted, using both **virtual** and **physical** platforms to ensure broad and effective communication. In the virtual realm, the project will use its **website** with a dedicated **blog** section to publish updates on the project's progress. This online resource will serve as a central hub for disseminating information and engaging with stakeholders interested in the project's developments. In addition, **social media platforms**, with a particular focus on **X (Twitter)**, will be used for direct communication. This approach allows for real-time interaction and provides a platform for stakeholders to engage with the project team, ask questions and provide

feedback. In addition, SoFAIR plans to organise **webinars** focusing on the results of the project and specific technological solutions that have been developed. These webinars will provide an interactive forum to present the results of the project to a wide audience and allow for in-depth discussion and clarification of technical aspects. Finally, the project will engage relevant stakeholders through relevant scientific **publications and presentations** during external events such as conferences, workshops and seminars. Additionally, a **mailing list** will also be prepared for individuals interested in the project outcomes, through which updates on progress will be shared.

3.2 Communication & Dissemination Activities

3.2.1 Internal communication

For internal team communication, mailing lists were set up for both the project as a whole and for individual work packages. In addition, a Slack channel was created to facilitate efficient communication and interactive discussions. Spreadsheets have also been developed to track the activities of project team members at national and international conferences and other meetings such as workshops and seminars. Moreover, regular meetings are held for the entire team as well as for individual smaller groups within the framework of work in work packages, enhancing the coordination and progress tracking of specific tasks and objectives.

3.2.2 External communication and dissemination

Website

The SoFAIR project places a strong emphasis on external communication through its online presence. A website has been developed to serve as a comprehensive source of information about the project, detailing its objectives, timeline of activities and the individuals and organisations involved.

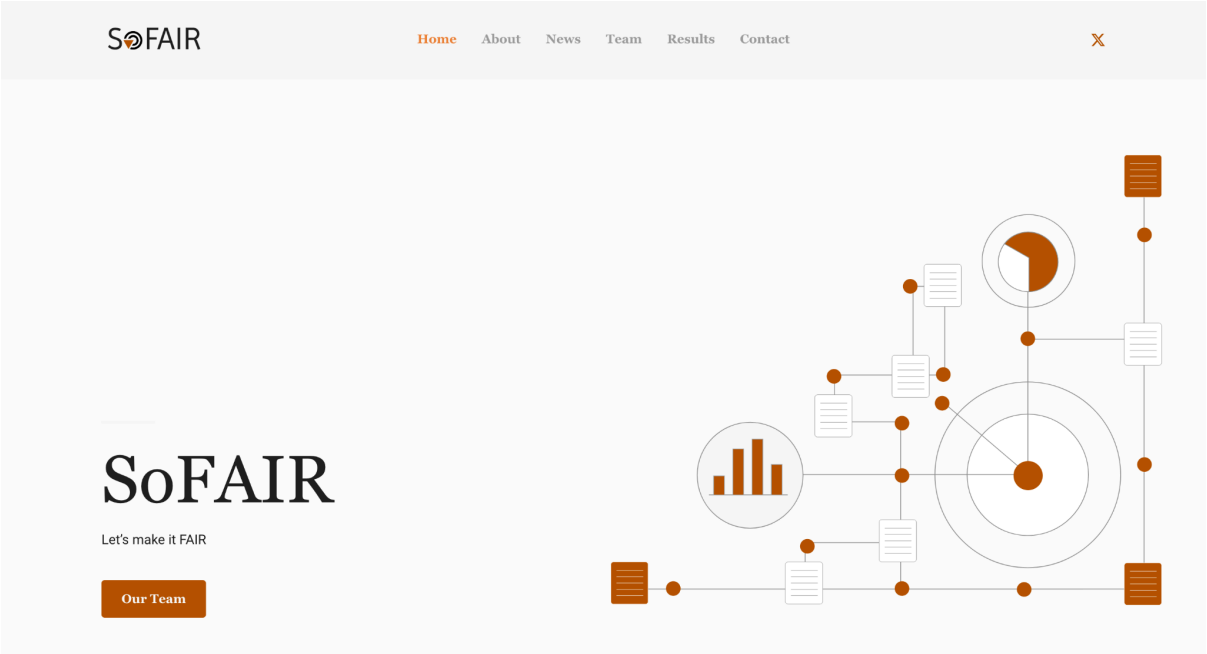

Fig. 1. Main page of project website

This platform serves not only to inform but also to engage with external stakeholders by providing updates on the project's progress and contributions to the field. The website includes a blog module, which is a key tool in the project's communication strategy. This module will be used to publish press releases, such as announcements about the launch of the project, milestones achieved and other important news. It will also include presentations by individual project partners and reports on the results of various activities.

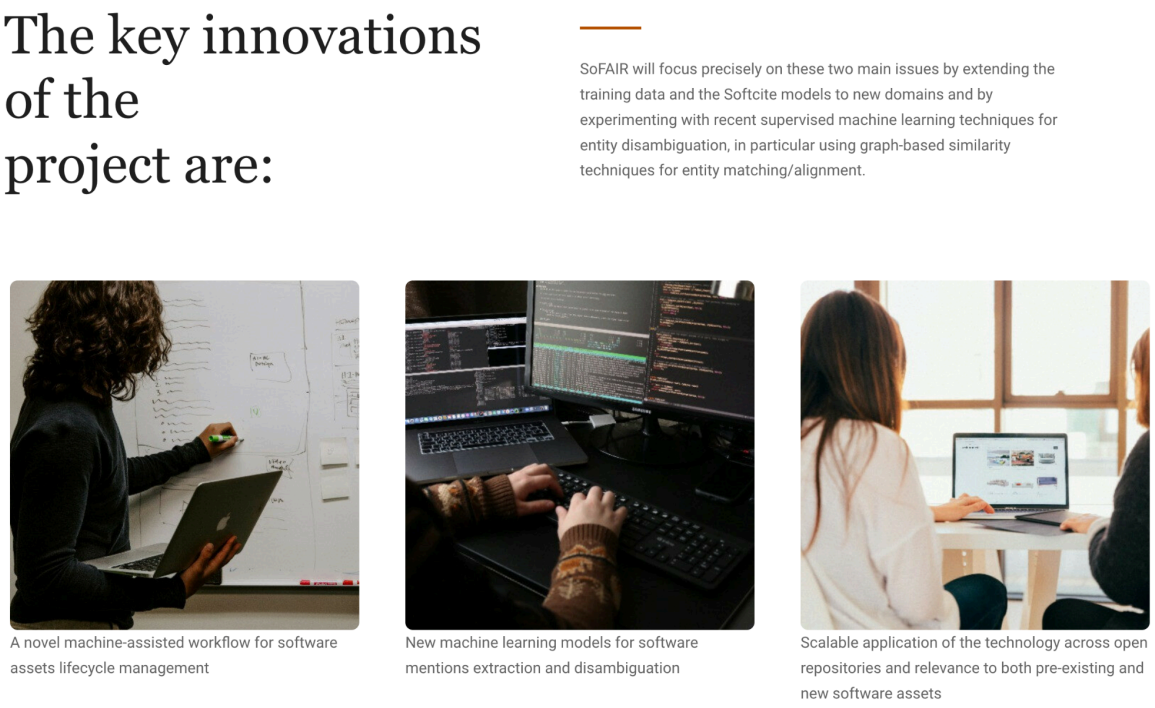

Fig. 2. Sample page content

The website features also a deliverables subpage and contact details to ensure easy access to further details and support, increasing the potential for collaboration and engagement.

Blog Posts plan

The blog will feature three categories of posts to meet the needs of SoFAIR's multiple stakeholder groups:

- **Partner presentations:** These blog posts will focus on introducing the project's partners, providing an insight into their roles, contributions and the expertise they bring to the project. This aspect of the blog aims to highlight the collaborative nature of the project and the diverse expertise involved.
- **Project updates:** These posts will cover key information about the project, providing a regular stream of news and updates to keep stakeholders informed of progress and developments.
- **Results and insights:** Posts in this category will focus on the specific outcomes of the project's activities. By detailing the achievements, innovations and insights gained through the project, these posts aim to showcase the tangible benefits and advances made possible by the SoFAIR project.

Social media – Twitter (X)

A Twitter (X) account has also been set up for the project, which will be used to publish short press releases, information about new blog entries and updates on the progress of the project. This account will be used to interact with individuals (including scientists, entrepreneurs, policymakers, etc.), institutions and other projects. We will also use the hashtag **#sofair** when tweeting. Stakeholder tweets will use both the hashtag and the project profile name.

Fig. 3. Projects Twitter (X) account

All publications, both on the website and on social media, will be planned in advance using internal tracking spreadsheets. These spreadsheets will include information such as the planned publication date, content, target audience and the person (or institution) responsible for creating the post.

Webinars

As a component of the project, a series of three webinars will be organised to disseminate the results of the project and to train participants in the solutions developed during the course of the project. These events will be aimed at a wide range of stakeholders who have expressed an interest in engaging with the project, from academic researchers and public institutions to individual practitioners. Each webinar is carefully tailored to the specific needs and skill levels of the audience, ensuring that the content is both accessible and relevant

External events participation

An integral part of the project is the active participation of project members in various conferences, workshops and seminars relevant to the thematic scope and objectives of the project. To optimise engagement and dissemination outcomes, a detailed tracking system will be used to catalogue all potentially relevant events, including comprehensive descriptions and detailed assessments of opportunities for presenting findings or publishing results. This proactive approach not only ensures the visibility of the project's contributions within the scientific community but also facilitates the exchange of knowledge and the maintenance of collaborative networks. Through strategic participation, the project aims to increase its impact and stimulate progress and innovation in the field.

Other publications

Within the scope of this project, a significant focus will be placed on the creation and dissemination of scientific publications, alongside the publication of preprints. This approach is designed to ensure that the research findings are shared with the scientific community in a timely manner, facilitating early feedback and fostering collaboration. The process of publishing preprints, in particular, enables the project team to make their research available to the public before the completion of the peer review process, thereby accelerating the dissemination of knowledge and findings.

3.4 Key Performance Indicators

Tab. 2. Key Performance Indicators

Activity	# at year 1	# at the end
Number of blog posts	5	10
Publications submitted (preprints available to the public)	0	2
Webinars	1	3
Number of tweets	50	100

In the first year of the project, the Key Performance Indicators (KPIs) are set to establish a solid foundation for communication and dissemination activities. The project plans to generate 5 blog posts, reflecting an initial effort to share insights, progress, and findings with a broader audience. At this stage, there are no publications submitted, indicating a focus on research and development activities. However, the project does aim to conduct 1 webinar, initiating direct engagement with stakeholders and disseminating early outcomes. Additionally, a goal of 50 tweets is targeted, showcasing an active engagement on social media to build a following and foster community interaction. By the end of the project, the KPIs are expected to show significant growth in all areas. The number of blog posts is projected to double, reaching a total of 10, demonstrating an increased capacity to share valuable content and insights. In terms of scholarly contributions, the project anticipates having 2 preprints available to the public, marking its achievements in producing tangible research outputs. The frequency of webinars is set to triple, with a total of 3 planned, enhancing opportunities for interactive learning and engagement with the project's findings. Lastly, the project aims to double its social media activity to 100 tweets, further expanding its reach and impact within the digital community.

4 Visual identification

Logotypes

Fig. 4. Main Logo

Fig. 5. Favicon logo in black

Fig. 6. Favicon logo in white

Colour palette

- #B75400 – Primary default (buttons colour)
- #EF8237 – Primary light
- #212121 – Black (Text colour)
- #9E9E9E – Dark grey
- #E0E0E0 – Medium grey
- #F5F5F5 – Light grey (background colour)

5 Risks/Threats

Tab. 3. Risk mitigation strategy

Risk	Description	Mitigation Strategy
Communication Balance	Difficulty in effectively communicating to diverse target audiences, leading to disengagement or misinterpretation.	Develop tailored communication strategies for each target group, incorporating feedback mechanisms to adjust strategies as needed.
Social Media Instability	The project relies on social media platforms which may experience downtime or changes in policy that affect outreach.	Diversify social media channels to reduce dependency on any single platform.
Mailing List Issues	Technical problems with mailing lists could impair direct communication with stakeholders.	Use multiple communication platforms (website, direct mail, slack channel) and maintain updated contact lists to ensure redundancy.
Internet Connectivity for Webinars	Live webinars depend on stable internet connections, which can be unpredictable.	Conduct technical rehearsals, have a recorded version ready for immediate release in case of live session failure, and choose reliable hosting services.
Website and Blog Instability	The project's website and blog may face downtime or technical issues, hindering information dissemination.	Implement regular website and blog maintenance schedules, use reliable hosting services, and prepare static versions of essential content for accessibility during downtimes.

List of Figures