Enhancing Trust, Integrity, and Efficiency in Research through Next-Level Reproducibility Impact Pathways

Deliverable D5.1 – Tools and practices for researchers

30/10/2025

Lead Beneficiary: ARC

Author/s: Eleni Adamidi, Thanasis Vergoulis, Fakhri Momeni, Elli Papadopoulou

Reviewer/s: Martin Reczko, Thomas Klebel

Funded by the European Union

Prepared under contract from the European Commission
Grant agreement No. 101094817

EU Horizon Europe Research and Innovation action

Project acronym:	**TIER2**
Project full title:	**Enhancing Trust, Integrity, and Efficiency in Research through Next-Level Reproducibility Impact Pathways**
Start of the project:	January 2023
Duration:	36 months
Project coordinator:	Dr. Tony Ross-Hellauer

Deliverable title:	Tools and practices for Researchers
Deliverable n°:	D5.1
Version n°:	1.0
Nature of the deliverable:	Report
Dissemination level:	Public

WP responsible:	WP5
Lead beneficiary:	ARC

TIER2 Project, Grant agreement No. 101094817

Due date of deliverable:	Month M34
Actual submission date:	Month n°

Deliverable status:

Version	Status	Date	Author(s)
1.0	Draft	03 September 2025	Eleni Adamidi, Thanasis Vergoulis (ARC)
1.1	Review	16 October 2025	Martin Reczko (Fleming), Thomas Klebel (KNOW)
1.1	Final	30 October 2025	Eleni Adamidi, Thanasis Vergoulis (ARC)

Table of contents

Executive Summary

Deliverable 5.1 (Tools and Practices for Researchers) presents the outputs developed under Task 5.1 of WP5 in the TIER2 project. It translates conceptual work from WPs 3 and 4 into practical solutions that help researchers integrate reproducibility into the daily conduct of their research. The deliverable introduces three tools, namely SCHEMA Lab, ARGOS, and Methods Hub as well as two practices, the Reproducibility Management Plan (RMP) and the Reproducibility Checklist. Together, these outputs form a toolkit that supports planning, documenting, executing, and sharing research in a transparent way.

SCHEMA Lab, developed by Athena RC, provides an open-source virtual laboratory for running containerized computational tasks and workflows and creating experiments. This tool enables researchers to design, execute, and monitor reproducible computational analyses within a friendly user interface but also programmatically.

ARGOS, led by OpenAIRE, extends the established DMP tool to support Reproducibility Management Plans (RMPs). By using machine-actionable templates aligned with the DMP Common Standard, ARGOS connects reproducibility planning to existing data-management and policy infrastructures.

The RMP practice complements ARGOS by defining the conceptual content and taxonomy of reproducibility planning. It provides funders and institutions with a structured framework and evaluation lens for assessing reproducibility requirements across domains.

The Reproducibility Checklist, coordinated by GESIS, offers a concise, validated instrument that helps researchers document essential methodological information, ensuring transparency and comparability of computational methods.

Methods Hub, also developed by GESIS, is a platform for depositing, documenting, and sharing computational methods based on the Reproducibility Checklist. It connects metadata, code, and interactive environments (e.g. Binder, Jupyter) to enable direct re-execution and verification.

Collectively, these tools and practices contribute to TIER2's broader goal of improving trust, integrity, and efficiency in research through actionable reproducibility infrastructures. They align with international standards (e.g. RO-Crate, GA4GH TES, DMP Common Standard), connect with policy and monitoring activities across pilots, and establish pathways for sustainability.

List of Abbreviations

EU – European Union
DMP – Data Management Plan
CWL – Common Workflow Language
RO-Crate – Research Object Crate
FAIR – Findable, Accessible, Interoperable and Reusable
KIPs – Key Impact Pathways

RMP – Reproducibility Management Plan
RPOs – Research Performing Organisations
SMP – Software Management Plan
UX – User Experience
WP – Work Package

1. Introduction

TIER2 addresses the challenge of poor reproducibility by developing solutions tailored to different research needs. Within this broader mission, Task 5.1 of WP5 focuses on researchers who require practical, accessible tools and practices to ensure reproducibility from the outset of their projects. It aims not only to provide technical solutions, but also to embed reproducibility as a routine part of planning, data collection, and analysis.

Deliverable 5.1 therefore introduces three tools and two practices: the SCHEMA lab tool, the ARGOS tool, the Methods Hub platform, the Reproducibility Management Plans (RMPs) and the Reproducibility checklist platform. Each has been co-designed with input from researchers in life, social, and computer sciences, ensuring that outputs are relevant across epistemic contexts.

By consolidating these activities, this deliverable lays the foundation for wider adoption of reproducibility practices and for strengthening trust in research results across disciplines.

2. Practical Tools and Practices for Researchers

This section presents the concrete tools and practices developed to support researchers in embedding reproducibility throughout their projects. Each subsection details one of the core outputs, the SCHEMA lab tool, the ARGOS tool, the Methods Hub platform, the Reproducibility Management Plans, and the Reproducibility checklist platform, highlighting their objectives, development activities, current progress, and next steps.

By structuring the section around these outputs, we provide a transparent view of how Task 5.1 has translated the scoping and co-creation work from WPs 3 and 4 into practical, discipline-sensitive solutions. Together, these tools form a toolkit that enables researchers to plan, document, and execute their work with reproducibility as a guiding principle.

2.1. SCHEMA lab tool

Related Pilot(s): Pilot 3
Responsible Organisation(s): ARC
Stakeholders Addressed: Researchers
Type of tool/ practice: software tool

2.1.1. Scope & objectives

Researchers across disciplines often struggle to execute and reproduce computational analyses due to fragmented infrastructures, lack of standardized execution environments, and insufficient capture of metadata and provenance. This creates barriers for transparency, validation, and reuse of computational results.

Before TIER2, the foundation for this work was the SCHeMa framework developed by Vergoulis et al. (2021), which introduced methods for scheduling and executing scientific containers on clusters of heterogeneous machines. This earlier work established the technical basis for orchestrating containerized computations but did not yet provide a user-facing interface or mechanisms for metadata capture, workflow management, or experiment packaging.

Building upon this research, TIER2 supported the conceptualization and development of SCHEMA api and SCHEMA lab (Adamidi et al., 2025), two new, open-source components designed to create scalable, user-friendly environments for reproducible computational research. Within TIER2, we implemented:

- **SCHEMA api**, which enables programmatic execution of containerized tasks and workflows while automatically recording metadata and provenance; and
- **SCHEMA lab**, a dedicated web-based interface that allows researchers to design and run computational experiments through an accessible virtual laboratory.

Together, these developments, entirely realized within the TIER2 framework, extend the earliest related work into a platform that embeds reproducibility, transparency, and ease of use into computational experimentation.

2.1.2. Development process & related activities

The development of SCHEMA lab and the underlying SCHEMA api followed an iterative process, ensuring that both the technical architecture and user experience were shaped by real research needs and community feedback. The process unfolded through four main phases:

- Design
  The initial design established the architectural foundation of the platform, focusing on a modular separation between the backend service (SCHEMA api) and the frontend interface (SCHEMA lab). The API was conceived to support reproducible, containerized execution of computational tasks in a standardized and interoperable manner, while the web interface was designed to make these capabilities accessible to researchers with varying levels of technical expertise. Early in this phase, the core operational concepts were defined: tasks, representing single executable units (e.g., a containerized software run), and workflows, representing linked tasks with defined dependencies.

- Co-creation
  To ensure that the platform reflected actual user needs, researchers from the Life Sciences Research ELIXIR communities were actively engaged in the co-creation process. Through questionnaires, online consultations, and interactive webinars, participants provided valuable input on usability aspects, metadata requirements, and the most common workflow use cases encountered in their research. This user-driven approach influenced interface design choices, metadata schema definition, and prioritization of features such as experiment grouping and monitoring dashboards.

- Testing
  The initial prototypes were internally deployed at Athena Research Center (ARC) for technical validation and user testing. Examples were tested using representative scientific workflows,

including the single-cell RNA-seq analysis workflow developed at the Biomedical Sciences Research Center "Alexander Fleming." The platform was also showcased in community events such as the ELIXIR 2024 BioHackathon and the SSDBM 2025 conference in Ohio.

- Refinement
  Based on continuous feedback from developers and early users, the system underwent several improvement cycles. Enhancements included the introduction of the "experiment" concept, enabling users to group multiple tasks or workflows under a single project entity with unified metadata capture; the development of a real-time monitoring dashboard for tracking execution progress and performance metrics; and the implementation of RO-Crate export, allowing complete computational experiments to be packaged for sharing and reuse.

Moreover, during the development process there was a strong alignment with WP4/WP5 activities, particularly through cross-pilot collaboration with Pilot 4 (Reproducibility Checklists) to explore the integration of reproducibility indicators into the creation of computational experiments.

2.1.3. Final Outcomes

During TIER2, a prototype version of SCHEMA lab was developed and released, featuring full integration with the SCHEMA api backend. Together, these components form an open-source framework that enables researchers to design, execute, and monitor computational experiments in a reproducible manner. We provide a list of outcome resources at the end of this section.

The system architecture illustrates how the SCHEMA lab front-end communicates with the SCHEMA api backend, which in turn interfaces with the Task Execution Engine (TESK, 2019), Kubernetes orchestration layer, and S3-based storage (see **Figure 1**). This modular architecture provides a scalable foundation for containerized computation.

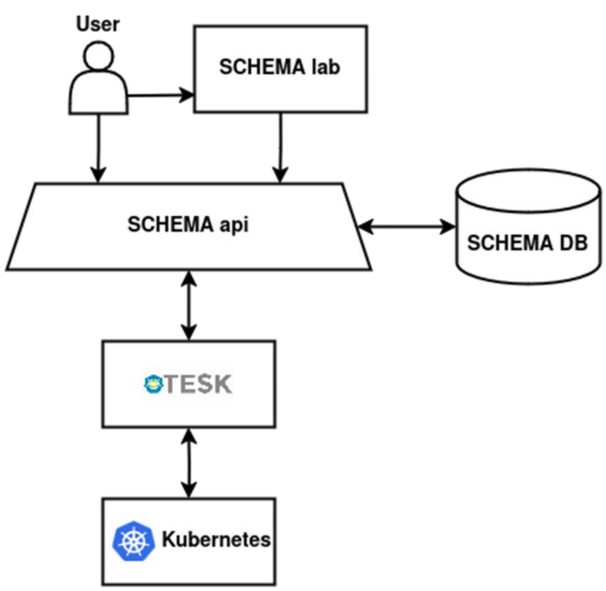

Figure 1: *High-level system architecture showing the interaction between SCHEMA lab, SCHEMA api, and the underlying task execution and storage layers (TESK, Kubernetes, S3).*

Features developed include:

- Execution of containerized software as tasks or workflows: Researchers can run containerized tools either as standalone tasks or as part of multi-step workflows, ensuring consistent and portable execution across environments.
- Creation of computational experiments with metadata capture: Tasks and workflows can be grouped into experiments, which capture execution metadata, parameters, and provenance information for transparency and reproducibility.
- Real-time monitoring of task and workflow executions: Users can track the progress and status of each computational run through the integrated dashboard.
- Export of experiments as RO-Crates: Completed experiments can be packaged and exported following the RO-Crate specification, enabling easy sharing and reuse across platforms.
- User interface for managing computational runs: SCHEMA lab provides an intuitive, dashboard-style web interface for creating, editing, and reviewing experiments and their associated executions.

A view of the main SCHEMA lab dashboard is shown in **Figure 2**, where users can monitor in real time the progress and status of all single tasks (Approved, Scheduled, Running, Cancelled, Rejected, Completed) through a visual, interactive interface.

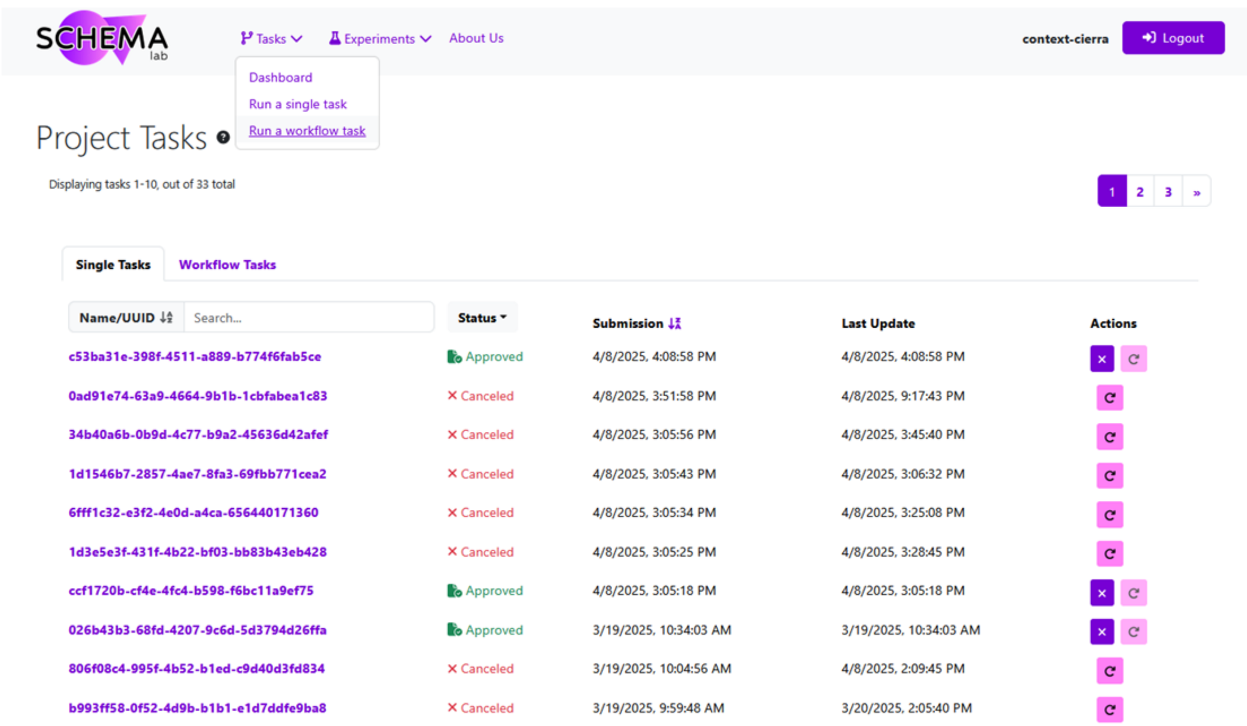

Figure 2: *SCHEMA lab dashboard displays real-time monitoring of single computational tasks (in the tab "Single Tasks).*

The following **Figure 3** presents the workflow dedicated dashboard for monitoring the status of workflows and perform actions such as cancel or rerun.

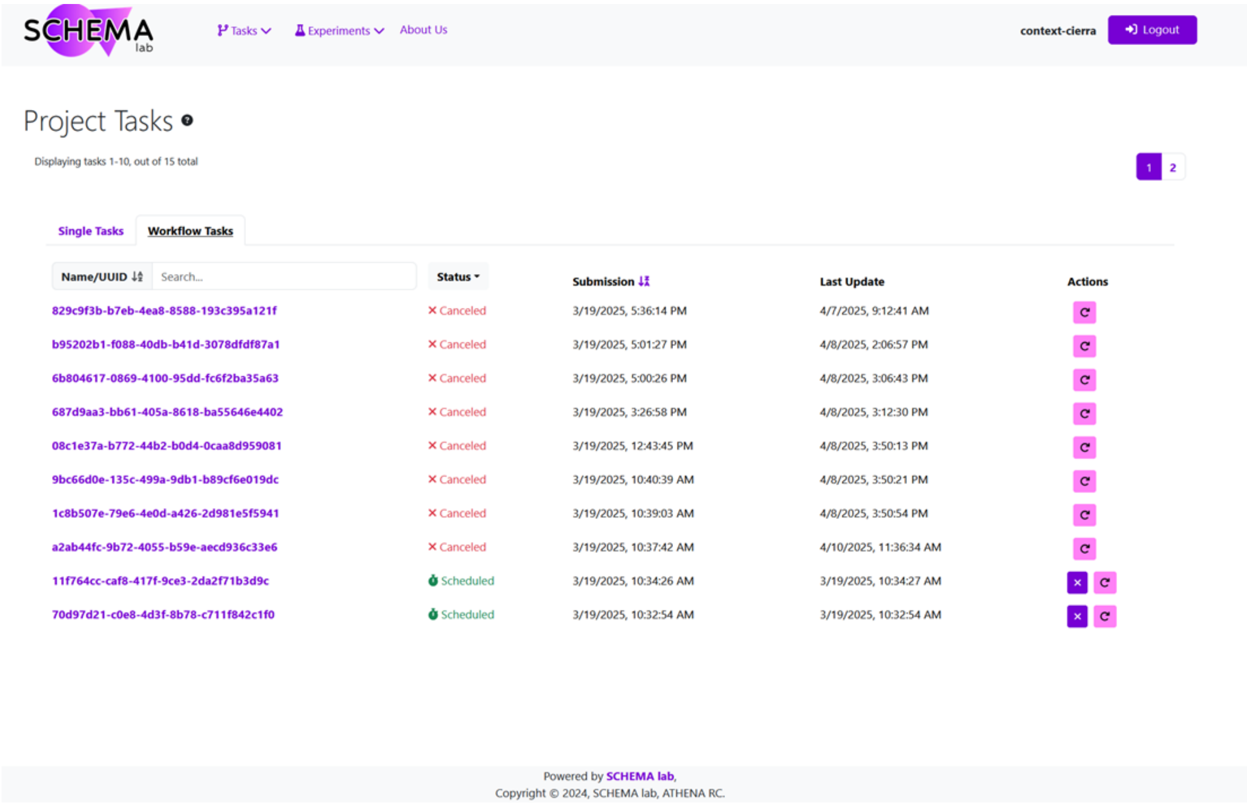

Figure 3: *SCHEMA lab dashboard displays real-time monitoring of computational workflows (in the tab "Workflow Tasks").*

The creation, management, and monitoring of computational experiments is presented in the dedicated dashboard in **Figure 4** where the user can select one or more successfully completed computational tasks or workflows and add them in an experiment.

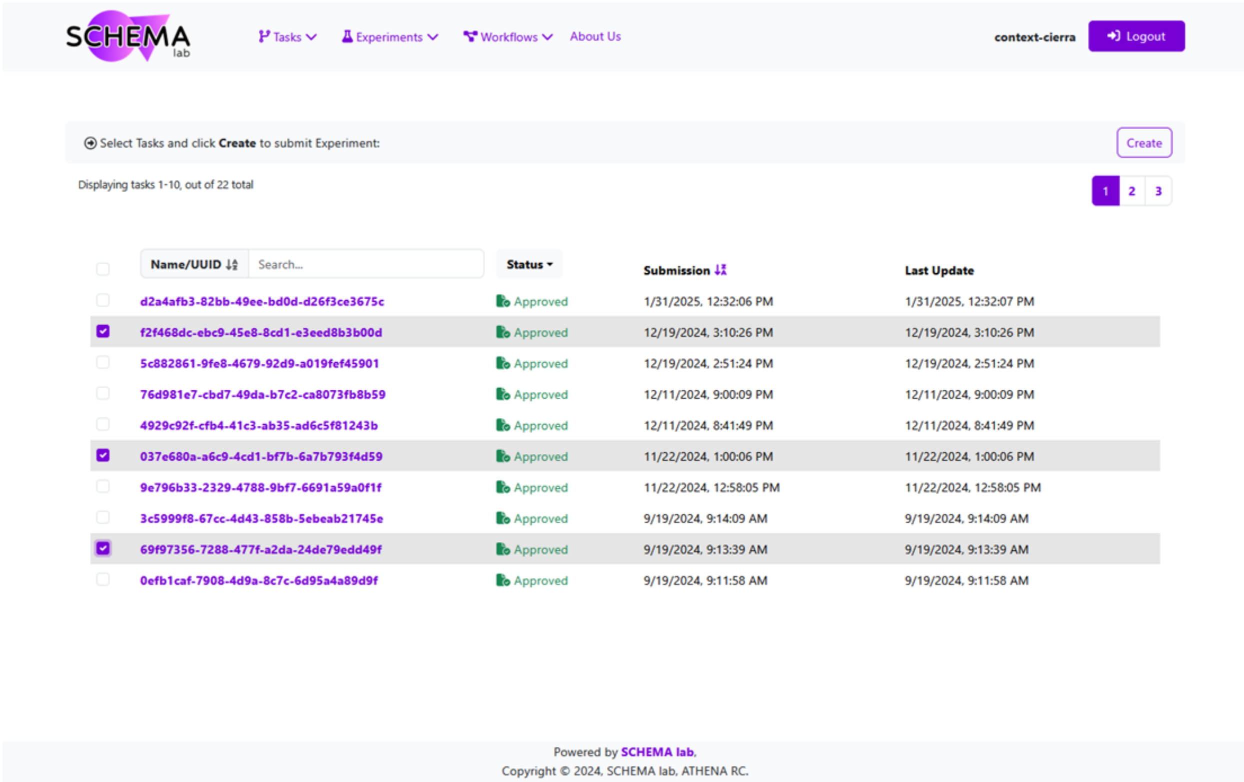

Figure 4: *SCHEMA lab dashboard displays real-time monitoring of computational workflows.*

The prototype and related resources are openly available through the following channels:

- SCHEMA lab homepage: https://schema-lab.hypatia-comp.athenarc.gr/
- GitHub repositories (open-source):
  - SCHEMA api: https://github.com/athenarc/schema
  - SCHEMA lab: https://github.com/athenarc/schema-lab
- ACM SSDBM 2025 paper: https://dl.acm.org/doi/10.1145/3733723.3733743
- Technical documentation (API specs, guides)
  - SCHEMA api swagger is available here: https://api.hypatia-comp.athenarc.gr/
- User documentation is available here: https://schema.athenarc.gr/

2.1.4. Value for Reproducibility

The SCHEMA lab and SCHEMA api together contribute significant value to the reproducibility of computational research by providing both the technical infrastructure and the monitoring capacity required for transparent, traceable, and repeatable analyses.

From a technical perspective, the platform offers a scalable environment for the reproducible execution of containerized computational tasks and workflows. By relying on containerization technologies (e.g., Docker images (Boettiger, 2015), k8s), SCHEMA ensures that analyses can be re-executed. Each run captures extensive information, including workflow, parameters, input and output datasets, image versions and execution settings, allowing researchers to reproduce computational experiments. Through its programmatic interface (SCHEMA api) and the user-friendly web application (SCHEMA lab), the platform enables researchers to design, execute, and monitor experiments in a consistent and well-documented manner. Furthermore, all completed experiments can be exported as RO-Crates, ensuring long-term accessibility and interoperability of research artefacts across systems and communities.

In terms of monitoring capacity, the platform automatically collects detailed metadata on each computational run, including resource utilization and time durations. This metadata forms a structured provenance record linking together the code, container image, input data, parameters, and resulting outputs. Such records make it possible to perform comparisons, enhancing the reliability of computational work. In addition, the dashboard interface provides a monitoring view of active and completed executions. This level of observability not only supports quality control but also enables researchers to evaluate reproducibility as a measurable attribute of research outputs.

The relevance of this tool spans multiple epistemic contexts. In the life sciences, SCHEMA lab has been applied to containerized pipelines such as single-cell RNA-seq analysis, providing transparent, repeatable execution. In the computer sciences, it supports the re-evaluation of experiments by allowing researchers to use RO-Crates, thus ensuring transparency across studies. Moreover, preliminary investigations are underway to extend its use to the social sciences, where scripted analyses (e.g., R or Python-based statistical workflows) can benefit from the same reproducible and provenance-aware execution environment.

In summary, SCHEMA lab and SCHEMA api extend beyond simple execution tools to function as a complete ecosystem for reproducible, transparent, and verifiable computational research. By embedding standards, metadata capture, and validation into everyday research practice, the platform helps make reproducibility a more practical and attainable aspect of computational work.

2.1.5. Stakeholder Engagement & Adoption

Stakeholder engagement was central to the development and early adoption of SCHEMA lab. Researchers from ELIXIR Greece and affiliated institutions were actively involved throughout the design process, providing feedback through questionnaires, webinars, and technical discussions that helped identify user needs, preferred interface features, and workflow requirements. This iterative engagement ensured that the platform addressed real challenges faced by researchers in executing and reproducing computational analyses. The prototype version of SCHEMA lab was run as a pilot service allowing selected users to experiment with running containerized tasks and workflows in a controlled environment. A concrete use case from the Biomedical Sciences Research Center “Alexander Fleming” was implemented to validate functionality and assess usability in a real research context. Beyond the project, SCHEMA lab has been demonstrated at international venues, including SSDBM 2025 and the ELIXIR 2024 BioHackathon, where it received positive feedback from both technical and scientific audiences. These early engagements demonstrate growing interest in adopting SCHEMA lab as a tool for reproducible computational research.

2.1.6. Sustainability & Future Use

The sustainability of SCHEMA lab beyond TIER2 relies on its strong alignment with community standards, its potential integration into existing research infrastructures, and its recognition within the European open science ecosystem. From the outset, SCHEMA lab and the underlying SCHEMA api were designed to build upon established interoperability standards, including the GA4GH Task Execution Service (TES) for orchestrating containerized computations and RO-

Crate for structured packaging of research artefacts. This standards-based foundation ensures that the platform remains compatible with broader international efforts to promote reproducible and transparent computational research.

Future development will focus on expanding the platform's interoperability and usability. Planned enhancements include support for widely used workflow description languages such as Nextflow, Snakemake, and Common Workflow Language (CWL), which will allow researchers to execute existing workflows seamlessly within SCHEMA lab. Further integration with WorkflowHub will enable workflows to be imported, executed, and shared in alignment with European workflow registries, while publication to RO-Hub will facilitate long-term preservation and discoverability of reproducible computational experiments.

Several key lessons have emerged during development and pilot testing. Researchers expressed that user-friendly interfaces that abstract technical complexity are essential to adoption, intuitive interaction and clear feedback mechanisms significantly lower the entry barrier for reproducibility practices. At the same time, while automatic metadata capture is a powerful enabler of reproducibility, it must be complemented by user-provided contextual information, such as research groups involved, funding or related to the specific research publications, to ensure the metadata remains meaningful. Those fields will be included in future versions to facilitate the creation of more complete experiments and RO-Crates.

2.2. Reproducibility Management Plan (RMP) practice

Related Pilot(s): Pilot 2
Responsible Organisation(s): OpenAIRE
Stakeholders Addressed: Researchers, Funders
Type of tool/ practice: contextual framework based on the practice of DMPs

2.2.1. Scope & objectives

The RMP practice addresses a gap in the theoretical approach of our work: there is currently no established 'RMP' construct so that RMP practices can flourish. We examine the full research lifecycle and introduce RMPs at the planning stage as a pragmatic way to overcome the issues we identify. Traditional DMPs focus on data and only partially cover the reproducibility pathway, and too often reproducibility is treated as an after-the-fact exercise pursued through reproducibility studies, rather than a proactive way to organise and connect research activities, roles, tools, and information. Our objective is to define the concept and content of an RMP: a structured, discipline-agnostic plan that articulates responsibilities, timelines, and the relationships between research outputs and the actions needed to reproduce or replicate findings. Where ARGOS provides the technical vehicle (templates, PIDs, exports), the RMP practice provides the content model and evaluation lens (question taxonomy, guidance, rubric).

2.2.2. Development process & related activities

We synthesised elements from DMPs, SMPs, datasheets, and domain protocols, together with TIER2's literature work, to derive a question taxonomy that consolidates known practices rather than inventing a new model, explicitly extending DMP practice to cover reproducibility. Through expert consultations and focus groups we co-defined core questions, grouped into families, e.g. design & registration, data, software. The Decision Aid Pilot was consulted to attach contextual guidance to the RMP concept. By design, the taxonomy separates declaration (what will be done) from implementation (how it is enacted). We then tested drafts with researchers and stewards. The result is a multi-layered prototype: a general core applicable across disciplines and optional modules for social, life, and computer sciences to address epistemic specificities.

2.2.3. Final Outcomes

The main outcomes of this development activity are:

- **Concept & content:** a formalised definition of the RMP and a domain-agnostic core question set with optional domain plugins. The overall structure of the RMP content model showing its domain-agnostic core and discipline-specific extensions is illustrated in **Figure 5**.

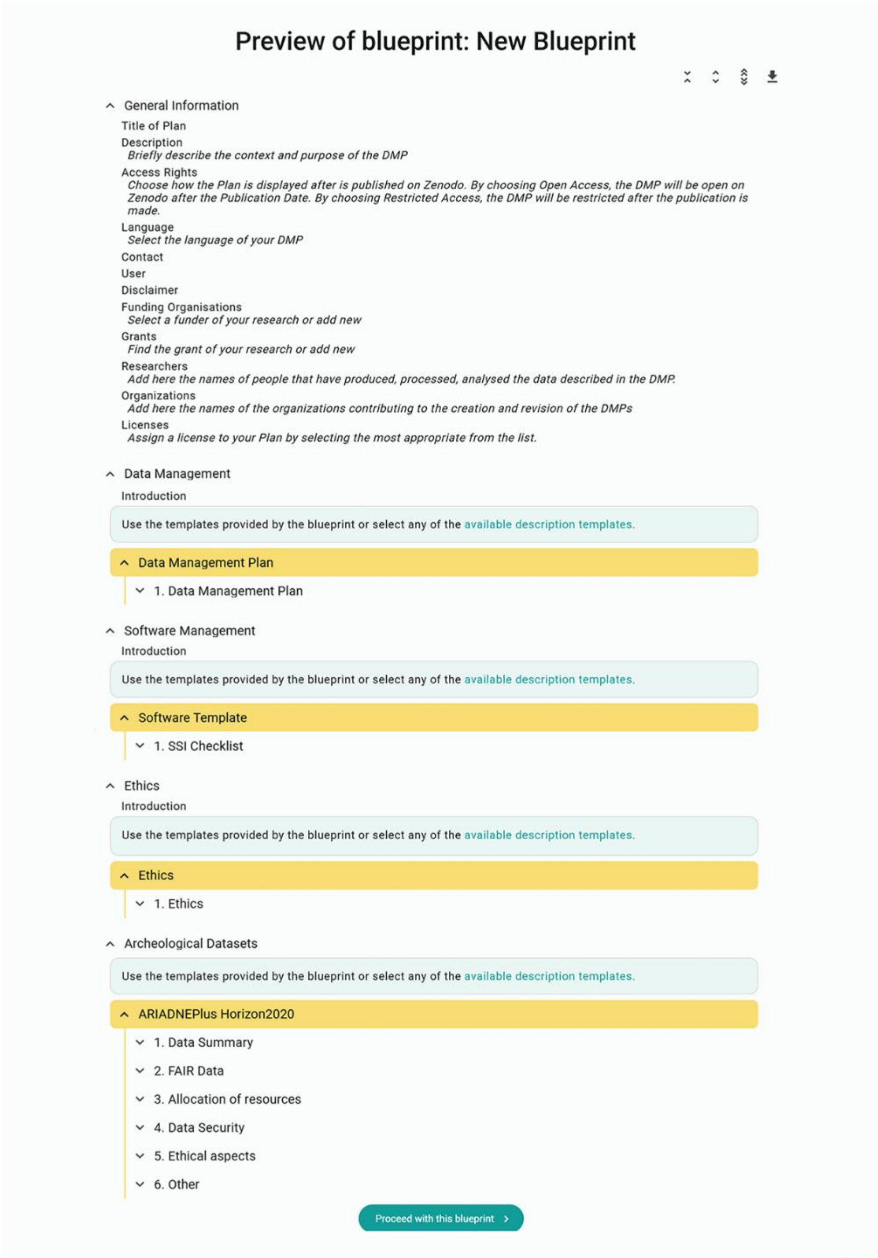

Figure 5: *Blueprint overview with core and domain-specific sections.*

- **Guidance:** a practical handbook for researchers with examples, available at OpenPlato: https://openplato.eu/

- **Integration and standardisation:** the RMP content model has been provided as input to configure the ARGOS tool (see Section 2.3), enabling machine-actionable implementation of RMP templates. In addition, proposed extensions to the RDA DMP Common Standard have been developed to incorporate reproducibility-relevant fields, facilitating interoperability and wider adoption across research infrastructures.

2.2.4. Value for Reproducibility

The RMP practice establishes a common content model and minimum information set for planning about reproducibility. Epistemic mapping ensures expectations are realistic yet comparable across domains. For funders and institutions, the RMP yields assessable criteria and a rubric that can be operationalized in reviews, audits, and monitoring, providing a consistent basis for policy alignment and programme-level comparisons.

2.2.5. Stakeholder Engagement & Adoption

Workshops, focus groups, and peer review (e.g., via RDA groups) shaped the content. Early adoption will be visible through pilot projects completing RMPs and through institutions adopting them.

2.2.6. Sustainability & Future Use

The practice is designed to outlive the project: it can be embedded in funder policies and institutional guidance, taught in training programmes, and maintained as a community-curated prototype. Alignment with the DMP Common Standard enables interoperability across tools; publication of RMPs as open outputs supports transparency and community learning. A light-weight governance model (versioning of the core and domain modules, documented change logs, and alignment to DMP-CS updates) keeps the practice stable while allowing evolution.

2.3. ARGOS tool

Related Pilot(s): Pilot 2
Responsible Organisation(s): OpenAIRE
Stakeholders Addressed: Researchers, Funders
Type of tool/ practice: Software tool with configurable templates, APIs, machine-actionable exports

2.3.1. Scope & objectives

ARGOS addresses a practical gap in planning for reproducibility. Originating as an established data management planning tool, it leverages machine-actionable capabilities to connect familiar, yet often disconnected, practices beyond Data Management Plans (DMPs) (e.g., Software Management Plans (SMPs), datasheets, preregistrations, and domain protocols). By reusing, combining, and extending these concepts to frame RMPs, it makes them computable at the planning stage. The aim is to technically enable the RMPs to be produced and shared as open, FAIR, machine-actionable artefacts. To our knowledge, no other platform unifies these practices at the planning stage. ARGOS does this by leveraging and extending DMP community standards and APIs (e.g., the DMP Common Standard) so that RMPs are machine-actionable; embedding open-science practices with persistent identifiers (DOI, ORCID, ROR) and qualified references to datasets, software, workflows, methods, and contributors; and exporting structured fields that dashboards and assessment tools can compute on.

2.3.2. Development process & related activities

Our work progressed through four phases with an emphasis on infrastructure:

1. **Design for interoperability:** requirements from TIER2 and literature translated into a field model that reuses the RDA DMP Common Standard, aligns with the OSTrails maDMP application profile and common APIs, defines an RMP profile/extension, and maps to PID ecosystems.

2. **Co-creation of admin assets:** funder/institutional template governance, roles/permissions, guidance strings, and validation rules were drafted with stakeholders (incl. RMP events, OSTrails collaborations and ARGOS community calls).

3. **Technical enablement:** onboarding of RMP templates; configuration of the FAIRsharing API and additional APIs for content enrichment and implementation of qualified references to enrich links with their policies, fields of science etc; export pipelines for machine-actionable content using DMP Common Standard JSON with OSTrails and RMP extensions; and connectors towards OpenAIRE Graph/MONITOR/CRIS/repositories.

4. **Testing & refinement:** usability testing improved form logic and navigation; iterations produced the publishable template set and admin documentation.

2.3.3. Final Outcomes

The ARGOS tool provides a complete, standards-based implementation environment for Reproducibility Management Plans (RMPs). The main outcomes are:

- **Data model & in-platform assets:** a formal RMP data model and machine-actionable template integrated in ARGOS (TRL 5), supporting structured planning and documentation of reproducibility.[1]

- **Interoperability:** machine-actionable exports (DMP Common Standard JSON with RMP extensions); FAIRsharing-backed lookups; PID support (ORCID, ROR, DOIs); and qualified references across datasets–publications–software–methods–workflows–

[1] https://github.com/OpenCDMP/OpenCDMP

contributors. The proposed extensions to the maDMP are shown in **Figure 6**, which highlights the new reproducibility-related fields (in green).

```
    { "dmp": {
"dmp_status": "draft",
        "related_identifiers": [{
            "identifier": "10.5281/zenodo.12345",
            "identifierType": "DOI",
            "relationType": "IsDocumentedBy",
        }],
        "policies": [{
            "title": "TUWien Open Science Policy",
            "description": "Describes institutional RDM
obligations",
            "policy_type": "institutional",
            "jurisdiction": "European Union",
            "related_identifiers": [{
                "identifier": "10.5281/zenodo.9876",
                "identifierType": "DOI",
                "relationType": "Describes",
            }]
        }],
        "dmp_license": {
            "license_ref":
"https://creativecommons.org/licenses/by/4.0/",
            "start_date": "2024-03-01",
            "access_rights": "open",
        },

        "title": "Horizon Europe Pilot - Plan"
    }
}
```

Figure 6: *Part of the proposed extensions to the maDMP (changes in green). Source:* [OSTrails - D4.2 Horizon Europe Report]().

- **Evaluation mechanism:** integration of the Decision Aid evaluator within the RMP template, allowing reproducibility criteria to be applied during both planning and review phases (see **Figure 7** for an overview of evaluation outputs).

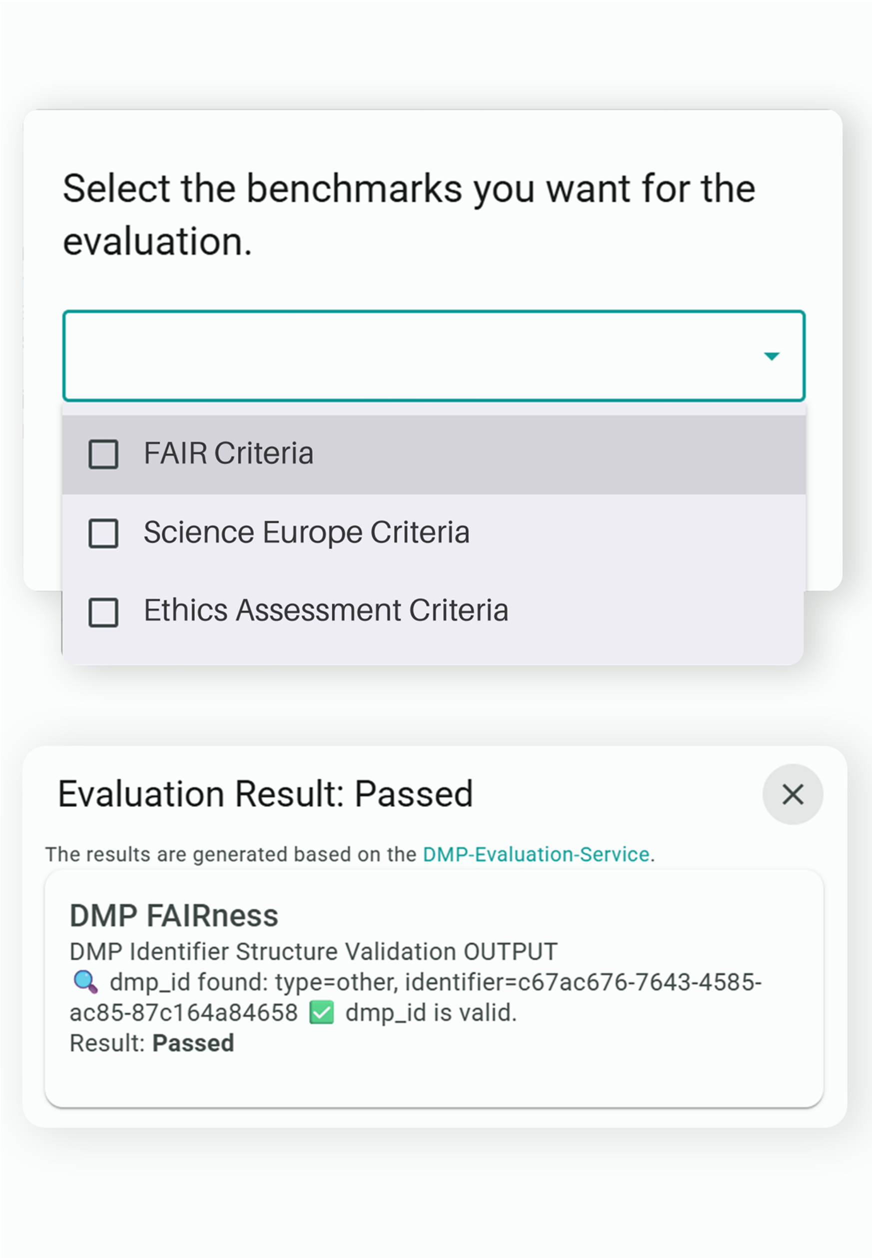

Figure 7: *Overview of OSTrails evaluator results.*

- **Operations:** step-by-step admin guides and configuration docs for funders/institutions; authoring guides for researchers and data stewards. All supporting materials, training resources, and documentation are openly accessible via the OpenPlato platform:
- ARGOS Service for Admins: https://openplato.eu/course/view.php?id=150
- ARGOS Info Pack: https://openplato.eu/course/view.php?id=547
- ARGOS Service for Users: https://openplato.eu/course/view.php?id=122

2.3.4. Value for Reproducibility

ARGOS provides a standards-based backbone for planning-stage reproducibility: an RMP data model aligned with the DMP Common Standard, PID-bound qualified references as indicated in the Horizon Europe framework programme, and export formats that downstream systems (e.g. monitoring dashboards, CRIS, repositories) can consume. In practice, RMPs become queryable records with auditable links to datasets, software, workflows, methods, and contributors, supporting policy indicators, automated checks, and evidence-based monitoring across programmes. The result is less manual reporting, less duplicate typing across reproducibility services (via API-driven auto-population of content to specific reproducibility activities), and clear, comparable visibility into practice at scale. Most importantly, it shifts reproducibility from retrospective checks to proactive, computable planning across people, tools, and activities.

2.3.5. Stakeholder Engagement & Adoption

We involved researchers and research project beneficiaries, data stewards, librarians and funders to shape the RMP prototype content, consolidating needs in a policy workshop with 17 funders that articulated requirements for machine-actionable outputs enabling systematic monitoring and defined minimum information requirements from funder perspectives. Monthly ARGOS community calls averaging 15 participants evolved from early requirements gathering through mid-pilot usability feedback to late-pilot advanced use cases, iterating the prototype design and functionality. Through this engagement process, the community identified 47 bugs/issues with 23 implemented as enhancements, crowd-sourced RMP content examples improving guidance quality, tested features enabling rapid iteration, built a community of practice, and fostered stakeholder connections.

Across all calls, 89 unique participants from 12 European and three non-European countries contributed, with 18 regular attendees demonstrating sustained community engagement (KPI 2). Stakeholder composition balanced researchers (58%), funders (19%), data stewards/librarians (15%), and others (8%), using multiple engagement methods. This broad input enhanced relevance and legitimacy while building a community of practice supporting sustained adoption. We successfully delivered a novel RMP concept not previously available, confirmed through literature review, formal definition in preparation for peer-reviewed publication, and recognition by the Advisory Board members (KPI 1). Eleven out of 17 funders reported reproducibility as an important indicator for DMP evaluation (KPI 4); however, the process is encouraged for adoption rather than mandated as in the case of data management.

Early adoption is evidenced by ## CHIST-ERA project RMPs completed to date and will continue to be tracked via configured deployments for projects and institutions and growth in published plans beyond DMPs.

2.3.6. Sustainability & Future Use

Post-TIER2, ARGOS will maintain the RMP capability, align mappings to evolving standards, and keep exports stable. The configuration model allows funders/institutions to version templates, add guidance, and more organically embed the RMP in calls and internal policies as the DMP ancestor.

2.4. Reproducibility Checklist tool

Related Pilot(s): Pilot 4
Responsible Organisation(s): GESIS
Stakeholders Addressed: Researchers
Type of tool/ practice: Checklist

2.4.1. Scope & objectives

The Reproducibility Checklist was developed to address the problem that researchers often provide incomplete or inconsistent documentation of their methods, which makes reproducibility difficult. The objective was to design a concise but comprehensive checklist that captures the essential information needed for reproducibility while remaining practical and easy to use. It provides clear guidance on what needs to be reported, supports standardization of method documentation for Computational Social Science, and serves as the structured metadata foundation for Methods Hub. Methods Hub is an online platform developed for storing, documenting, and sharing computational methods in a standardized and sustainable way (presented in the next section 2.7 of this report. By lowering the entry barrier for researchers, the checklist makes reproducibility a more feasible and achievable part of everyday research practice.

2.4.2. Development process & related activities

The checklist was designed through an iterative co-creation process. Initial drafts were informed by state-of-the-art reviews and refined through two surveys with researchers in Computational Social Science. These surveys provided feedback on relevance, feasibility, and usability, ensuring the checklist balanced completeness with simplicity. An experiment study then tested how well researchers could reproduce methods using the checklist, offering systematic insights into its effectiveness. Finally, the checklist was technically integrated into Methods Hub, aligning the practice with the platform infrastructure.

2.4.3. Final Outcomes

The outcome of this work is a validated and simplified Reproducibility Checklist integrated into the Methods Hub ecosystem. The checklist captures the essential elements researchers must document to make their work reproducible, including information about data, code, computational environment, and dependencies. It was refined through two community surveys and tested in an experimental evaluation of reproducibility outcomes. Based on user feedback, the initial three-phase, detailed checklist developed from survey results was streamlined into a concise, easy-to-apply version to encourage adoption and minimize user effort. This simplified version focuses on core reproducibility requirements and is directly implemented in the Methods Hub submission workflow as structured metadata, ensuring that reproducibility information is included in every published method entry. In this way, the checklist has become an operational component of the platform rather than a separate document, supporting reproducibility in real research contexts.

The checklist template resources are publicly available through the **Methods Hub GitHub repository**:

- **README Template: Specification and Illustration of the Methods Hub Friendly README** – defines the structure and mandatory documentation fields that implement the reproducibility checklist items (e.g., environment setup, input/output data, code dependencies, and technical documentation).

- **Method Submission Guidelines** – provide user guidance and documentation quality criteria explaining how to apply the checklist when preparing and submitting methods.

2.4.4. Value for Reproducibility

The value of the reproducibility checklist lies in making computational methods more understandable and easier *to use* for social scientists who work with digital behavioral data or text analysis. By requiring only the essential information, the checklist lowers technical barriers and enables researchers with a *basic programming background* to reproduce and apply methods. This not only supports transparency and comparability within computational social science but also broadens access, ensuring that methods can be reused beyond highly technical expert communities.

2.4.5. Stakeholder Engagement & Adoption

The reproducibility checklist was shaped through active engagement with researchers in computational social science. Two surveys collected feedback on relevance, feasibility, and usability, ensuring that the checklist reflected real needs rather than abstract requirements. In addition, the checklist was tested in an *experiment study*, where researchers attempted to reproduce methods using it. This provided systematic evidence of its strengths and areas for refinement. Early signs of adoption are visible through its *integration into Methods Hub*, where it now serves as the metadata template for submitting methods.

2.4.6. Sustainability & Future Use

The reproducibility checklist is designed to continue as a practical standard within Methods Hub, where it functions as the metadata template for method submissions. Beyond the project, it can be integrated into existing infrastructures such as institutional repositories or national research data services, and it aligns with wider standards for open and reproducible science. Its simplicity also makes it adaptable for use by funders and publishers as a requirement for method documentation.

A key lesson learned is that checklists must remain short and practical to ensure adoption. Researchers are more likely to use a tool that minimizes additional workload while still capturing essential reproducibility information. Future development should therefore focus on refining usability, providing discipline-specific examples, and exploring interoperability with other reproducibility frameworks and infrastructures.

2.5. Methods Hub

Related Pilot(s): Pilot 4
Responsible Organisation(s): GESIS
Stakeholders Addressed: Researchers
Type of tool/ practice: Software tool

2.5.1. Scope & objectives

Methods Hub is an existing online platform developed at GESIS to store, document, and share computational methods in a standardized and sustainable way. It addresses the lack of dedicated infrastructures where researchers can deposit methods with sufficient metadata to ensure reproducibility and reuse. Within the TIER2 project, the platform served as the implementation environment for Pilot 4, focusing on enhancing the reproducibility and usability of hosted methods. The pilot contributed by developing and integrating a Reproducibility Checklist, conducting user studies to assess usability and content structure, and performing an experimental evaluation to measure the reproducibility of methods published on the platform. Through these contributions, TIER2 strengthened the reproducibility layer of Methods Hub, ensuring that deposited methods meet minimal reproducibility standards while keeping the process simple and researcher friendly. The platform's overarching goal remains to provide a long-term infrastructure that lowers barriers to sharing methods and enables reproducibility at scale.

2.5.2. Development process & related activities

The development of Methods Hub followed a co-creation and testing approach closely linked to the checklist work. The platform was designed in parallel with the checklist to ensure seamless integration. User workshops and expert interviews at GESIS informed early design choices, focusing on metadata requirements and usability of the front end. Iterative prototypes were tested with researchers during pilot activities and in the experiment study, with feedback leading to refinements in metadata schema, user interface, and integration with Jupyter/Binder environments. Through this process, Methods Hub evolved into a functioning beta platform grounded in researcher needs and reproducibility requirements.

2.5.3. Final Outcomes

The main outcome of this activity is a beta release of the Methods Hub platform, providing a functional environment for depositing, documenting, and sharing computational methods. The platform integrates the reproducibility checklist as structured metadata, ensuring that each method submission captures the minimum information required for reuse.

The beta version includes a working front end, metadata schema, and integration with Jupyter and Binder, allowing users to test and run methods directly. Methods Hub is accessible via methodshub.gesis.org and is accompanied by user documentation and developer guidance to support adoption by researchers and technical teams.

Figure 7.1 shows a Methods Hub interface displaying a list of available methods. By selecting a specific method, users can access a detailed view that presents all checklist information associated with that method (**Figure 7.2**). These structured fields make each method transparent, understandable, and easy to reproduce, even for researchers with basic programming experience. The detailed page also includes linkable metadata, such as direct links to interactive environments (e.g., mybinder) that allow users to reproduce the method step by step.

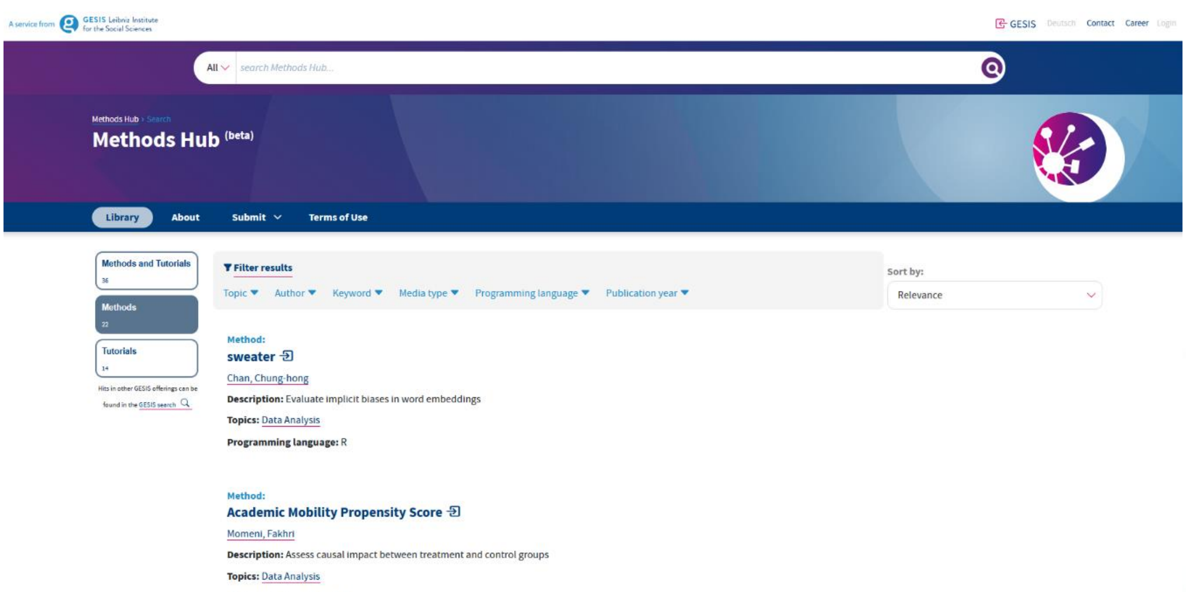

Figure 7.1: *Methods Hub overview page with a list of available methods.*

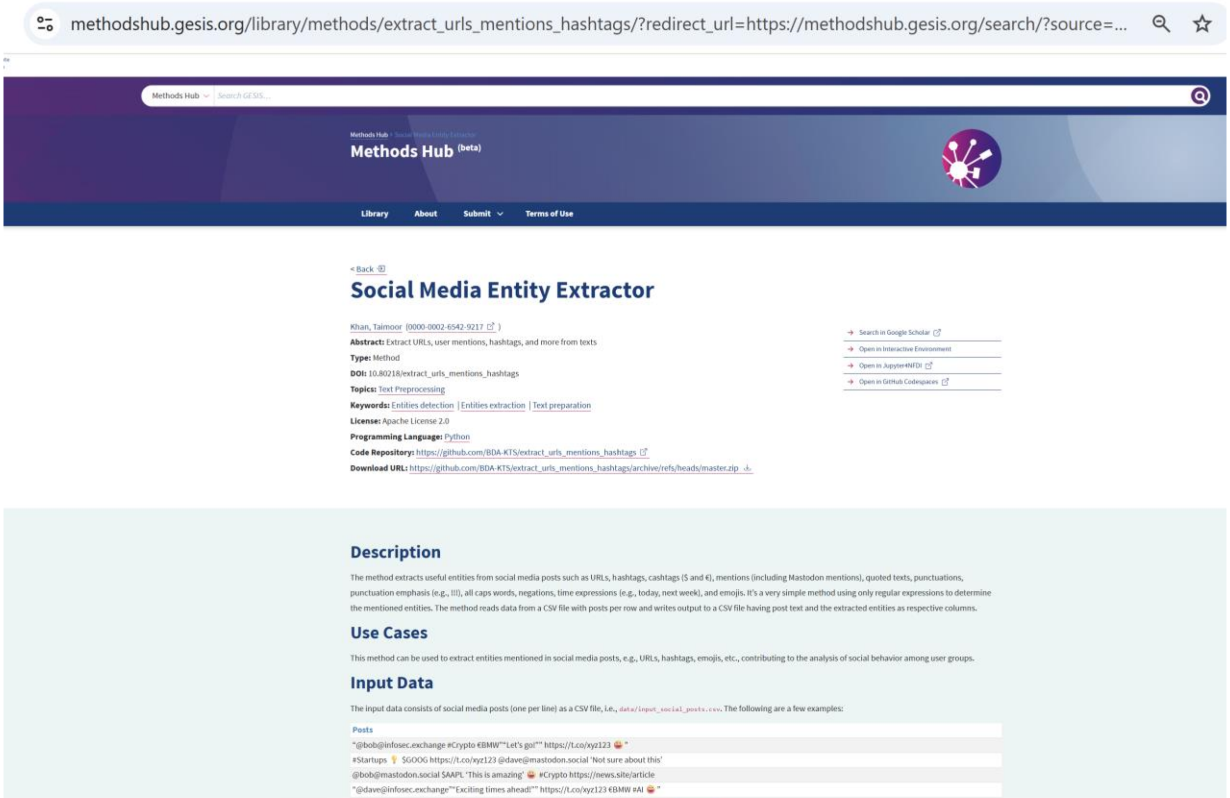

Figure 7.2: *Detailed method view in Methods Hub displaying the reproducibility checklist fields, metadata, and interactive links for reproducing the method (e.g., via Binder).*

2.5.4. Value for Reproducibility

Methods Hub contributes to reproducibility by offering a dedicated technical infrastructure for depositing and sharing computational methods. By integrating the reproducibility checklist as metadata, it ensures that every method is accompanied by the essential information needed for reuse and transparency. The platform enables not only the storage and documentation of methods but also their execution and testing through Jupyter and Binder, which makes reproducibility directly verifiable.

Although developed primarily for computational social science, Methods Hub provides a generic framework that can be extended to other fields such as computer science or life sciences. It also creates value for publishers and funders, who can use the platform as evidence of method documentation and reproducibility in line with open science and FAIR principles. In this way, Methods Hub supports both community-driven research practices and policy-level adoption of reproducibility standards.

2.5.5. Stakeholder Engagement & Adoption

The development of Methods Hub was guided by continuous engagement with researchers and experts. Early prototypes were presented in usability test and expert interviews at GESIS, focusing on the metadata requirements, the usibility and content. Feedback from these sessions led to simplifications in the submission process and improvements to the user interface.

In addition, Methods Hub was evaluated through an experiment study, where researchers tested the platform by reproducing methods. This provided systematic evidence of how well the platform supports usability and reproducibility in practice.

In addition to pilots and usability tests, Methods Hub was presented and discussed in a dedicated hands-on session at the MethodsNet conference 2025 and International Workshop on Computational Reproducibility in Social Sciences at ICWSM conference 2025, where researchers had the opportunity to try out the platform and provide direct feedback. This broadened engagement beyond the project consortium and created early visibility for the tool in the wider research community.

2.5.6. Sustainability & Future Use

Methods Hub is designed as a sustainable infrastructure that will continue beyond the TIER2 project. The platform is hosted and maintained at GESIS, ensuring long-term accessibility for researchers in computational social science and related fields. It is already integrated into *GESIS Search*, which allows methods deposited in Methods Hub to be indexed and discovered alongside publications and datasets. This positions Methods Hub within established research infrastructures and aligns it with open science and FAIR principles.

The platform also offers potential for policy-level adoption: funders and publishers could use Methods Hub as a reference point for verifying whether computational methods are sufficiently documented and reproducible. In this way, Methods Hub can serve both community-driven research practices and broader reproducibility requirements.

Key lessons learned during development highlight that usability is critical for adoption. Researchers are more likely to submit methods when the process is simple, streamlined, and requires only essential metadata. Future development should therefore focus on enhancing user experience, providing discipline-specific templates, and building interoperability with other reproducibility frameworks and infrastructures to increase uptake and scalability.

3. Synthesis

Taken together, Deliverable 5.1 has resulted in three tools (SCHEMA lab, ARGOS, Methods Hub) and two practices (Reproducibility Management Plans (RMPs) and Reproducibility Checklists). While distinct in scope and disciplinary focus, these outputs share a common purpose: to enable researchers to plan, document, and execute their studies with reproducibility as a core principle.

Collectively, these developments represent a concrete step forward in operationalizing the TIER2 vision, moving from conceptual frameworks and policy recommendations to practical, usable instruments that embed reproducibility in the everyday workflow of researchers. Each tool addresses a different stage of the research lifecycle: ARGOS and RMPs strengthen reproducibility at the planning stage, SCHEMA lab supports execution and management of computational tasks, and Methods Hub together with the Reproducibility Checklist ensure sharing, validation, and reuse of computational methods. This alignment ensures continuity across stages of research, turning reproducibility from a one-off activity into a sustained, integrated process.

Furthermore, the outputs developed under Task 5.1 are aligned with widely recognized community standards (e.g., RO-Crate, GA4GH TES, DMP Common Standard), ensuring interoperability with the broader open science landscape. This standards-based approach positions the tools for long-term sustainability and potential integration within infrastructures such as ELIXIR and OpenAIRE but also the broader research community.

From a broader perspective, Deliverable 5.1 demonstrates how technical tools and reproducibility practices can complement and reinforce each other. Tools such as SCHEMA lab and Methods Hub make reproducibility tangible and verifiable, while practices such as RMPs and Checklists provide the conceptual and procedural scaffolding that allows these tools to be used effectively. This interplay between infrastructure and practice is central to the TIER2 mission so that reproducibility is enabled as an integral element of trustworthy research.

In sum, the deliverable provides a coherent toolkit that supports transparency and accountability in research. It lays the groundwork for the continuation of TIER2's efforts beyond the project's lifetime, through open-source development, standard alignment, and integration into enduring research infrastructures, ensuring that reproducibility becomes a sustained and measurable part of the scientific process.

4. References

Vergoulis, T., Zagganas, K., Kavouras, L., Reczko, M., Sartzetakis, S., & Dalamagas, T. (2021). *SCHeMa: Scheduling scientific containers on a cluster of heterogeneous machines.* In *33rd International Conference on Scientific and Statistical Database Management (SSDBM 2021)* (pp. 243–247). Association for Computing Machinery. https://doi.org/10.1145/3468791.3468813

Adamidi, E., Deligiannis, P., Foutris, N., & Vergoulis, T. (2025). *A virtual laboratory for managing computational experiments*. In *Proceedings of the 37th International Conference on Scalable Scientific Data Management (SSDBM '25)* (Article 14, pp. 1–6). Association for Computing Machinery. https://doi.org/10.1145/3733723.3733743

Boettiger, C. (2015). An introduction to Docker for reproducible research. *ACM SIGOPS Operating Systems Review, 49*(1), 71–79. https://doi.org/10.1145/2723872.2723882

TESK: Task Execution Service for Kubernetes, 2019. https://github.com/EMBL-EBI-TSI/tesk