# STREAMING GIRLS AND WOMEN INTO STEAM EDUCATION, INNOVATION AND RESEARCH

**STREAM IT**

## D1.3 Data Management Plan

30.06.2024

STREAM IT

Project full title

# STREAMING GIRLS AND WOMEN INTO STEAM EDUCATION, INNOVATION AND RESEARCH

Project acronym

# STREAM IT

Project Agreement no.

**101131843**

D1.3 Data Management Plan

*Copyright by the STREAM IT consortium*

**DOCUMENT CONTROL SHEET**

| | |
|---|---|
| **AUTHOR(S)** | Dhanesh Chathuranga, Jan Ertmann |
| **REVIEWER(S)** | Réka Geambașu, Katalin Oborni |
| **DISSEMINATION LEVEL[1]** | PU |
| **START DATE OF PROJECT** | 01.01.2024 |
| **DURATION OF THE PROJECT** | 31.12.2026 |
| **SUBMISSION DATE** | 30.06.2024 |
| **ORGANISATION NAME OF LEAD CONTRACTOR FOR THIS DELIVERABLE** | S2i |

PU – Public (fully open, automatically posted online on the Project Result platforms);
SE – Sensitive (limited under the conditions of the Grant Agreement);
CO – EU classified : EU restricted, EU confidential, EU secret under Decision 2015/444

**VERSIONING AND CONTRIBUTION HISTORY**

| VERSION | DATE | MODIFICATION REASON | MODIFIED BY |
|---|---|---|---|
| 1.0 | 27.06.2024 | First draft | Dhanesh Chathuranga, Jan Ertmann (S2i) |
| 1.0 | 29.06.2024 | Quality Assurance checking | Réka Geambașu (UBB), Katalin Oborni (HETFA) |
| 1.0 | 30.06.2024 | Initial version of the Data Management Plan | Dhanesh Chathuranga, Jan Ertmann (S2i) |

# TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF ABBREVIATIONS

| DMP   | Data Management Plan                                      |
|-------|----------------------------------------------------------|
| FAIR  | Findable, Accessible, Interoperable, and Reusable        |
| HE    | Horizon Europe                                           |
| IP    | Intellectual Property                                    |
| QAP   | Quality Assurance Plan                                   |
| STEM  | Science, Technology, Engineering and Mathematics         |
| STEAM | Science, Technology, Engineering, Arts and Mathematics   |
| WP    | Work Package                                             |

## EXECUTIVE SUMMARY

ST(R)E(A)M IT is a pioneering initiative aimed at catalysing change within the persistent gender disparities in STEM education, research, and innovation. Aligned with the objectives outlined in 'The European Manifesto for gender-inclusive STE(A)M education and careers,' the project seeks to empower underrepresented groups, integrate STEAM methodologies into STEM education, promote gender-inclusive career pathways, and support STEM education providers in dismantling gender-based barriers.

Integral to the success of ST(R)E(A)M IT is the implementation of a robust Data Management Plan (DMP), Deliverable D.1.3. This DMP outlines the strategies for managing data within the project and underscores the crucial role of data in advancing the project's objectives. By adhering to principles of FAIR data (Findable, Accessible, Interoperable, and Reusable), the DMP delineates procedures for handling, sharing, and preserving research data throughout and beyond the project's duration.

Key components of the DMP include protocols for data handling during and post-project, identification of data types, application of methodologies and standards, provision for open access to data, and strategies for data curation and preservation. Moreover, the DMP articulates the project's policy on scientific publications, emphasising open access and leveraging Zenodo as the primary research data repository. Importantly, ethical considerations and data security measures are thoroughly addressed to ensure responsible data management practices.

Through the collaborative efforts of STEM-oriented entities, including businesses, educational institutions, research organisations, and civil society organisations, ST(R)E(A)M IT endeavours to bridge the gender gap in STEM education and careers. The DMP serves as a dynamic document, evolving in tandem with the project's progress, and will be periodically updated and reviewed to ensure its relevance and effectiveness.

# 1. INTRODUCTION

Data Management Plans (DMPs) represent a fundamental aspect of effective data management practices and are obligatory for all Horizon Europe projects. These plans outline the lifecycle of data to be collected, processed, or generated throughout the project duration. In alignment with the principles of making research data Findable, Accessible, Interoperable, and Reusable (FAIR), a comprehensive DMP encompasses various aspects, including data handling during and post-project, data types, methodologies and standards, data sharing through open access, and strategies for data curation and preservation.

## 1.1 PURPOSE:

Horizon Europe has progressed towards embracing open science, as outlined in the Horizon Europe Regulation, which establishes the legal framework for open science obligations and incentives applicable to Horizon Europe beneficiaries. Open science advocates for collaborative work, tools, and knowledge diffusion, encompassing open access to scientific publications and research data. Mandates for open access to both publications and research data underscore the commitment to openness, guided by the principle of 'as open as possible, as closed as necessary.'

This DMP serves as a dynamic document, subject to updates throughout the project period to address significant changes such as new data types, alterations in consortium policies, or shifts in consortium composition. The initial version of the DMP provides an overview of the envisaged data sets to be generated by the STREAM IT project. Steinbeis is entrusted with outlining the data management responsibilities.

## 1.2 PARTNER CONTRIBUTIONS:

The formulation of this initial version of the DMP was a collaborative effort involving all partners, ensuring alignment with the consortium's exploitation and Intellectual Property (IP) requirements. Adhering to the FAIR principles outlined in the STREAM IT Grant Agreement, beneficiaries are obliged to responsibly manage the digital research data generated within the project. Partners actively participated in providing input through a xls survey, sharing insights into their key data sets and planned data management strategies. Steinbeis contributed additional questions to elaborate on data management aspects, fostering comprehensive planning and implementation across the consortium.

## 2. Purpose of Data Collection

All data collected in the STREAM IT project aims to further the cause of gender inclusivity in STEM using STEAM approaches. To achieve these goals, STREAM IT will collect and produce data for both internal and external use. This includes analyses, reports, plans, case studies, surveys, interviews, workshops, online networking, webinars, a virtual makerspace, policy dialogues and event registrations. Datasets containing or derived from personal data will be explicitly specified to implement special protective measures, including anonymisation procedures. Special attention is given to the Nurturing Talent - Worksop as this is aimed at girls between 14-18 years. It will only allow participants where consent forms from the parents and guardians have already been signed; the results are only used for impact assessment and not published with personal identifiable information.

The table below outlines the essential aspects of data management for the STREAM IT project, ensuring compliance with GDPR and the efficient handling of various data types. The "Type of Data" column specifies the nature of the data being collected, such as research results or participant information. "File Format(s)" indicates the formats in which data will be stored, for example, PDF, DOCX, or CSV. "Approval for Planned Usage" details the necessary consents and permissions obtained for using the data. "How Data Will Be Collected" describes the methods employed for data collection, such as surveys, interviews, or automated systems, alongside the "Terms of Use" that participants agree to. The "Purpose of Processing" column explains why the data is being collected and how it will be used to support the project's goals. "Storage Location" specifies where the data will be securely stored, whether in institutional repositories, cloud services, or local servers. Finally, "Who Will Have Access to the Data" lists the authorised partners or groups that will be granted access to the data, ensuring that only relevant and approved individuals can handle sensitive information.

## STREAM IT

| Type of data | File format(s) | Approval for planned usage | How data will be collected (terms of use) | Purpose of processing | Storage Location | Who will have access to the data |
|---|---|---|---|---|---|---|
| **For all Persons involved** | Ethical considerations and GDPR applies | | | | | |
| **Nurturing Talent - Worksop series for supporting girls (14-18 years) to orient towards STEM** | Forms, Json, xlsx | Consent forms for parents and guardians and participants | There will be anonymous pre and post surveys and participation registration all of which will be covered by a consent form signed by the parents and/or legal guardians | Evaluating the educational activities and Impact assessment to finalise the toolkits in WP3 | Gdrive STREAM IT | Only involved STREAM IT Partners and the HETFA |
| **Webinars and training registration** | Basecamp, Forms, xlsx, | Registration and acceptance of privacy policy of HETFA | Apart from registration there is no foreseen data that will collected | Accomplishing KPIs and quantitative metrics | Gdrive STREAM IT | Only involved STREAM IT Partners and the HETFA |
| **Networking tool under virtual Makerspace** | None specified | Registration and acceptance of privacy policy of HETFA | Data will not be collected. | Registration is used to participate in the online forum. | STREAM IT Website | Only involved STREAM IT Partners and the HETFA |
| **Project website and social media datasets** | .docx, .pdf Photos & images (i.e. .jpg), | Consent forms of all persons contributing and sharing success stories | Through questionnaire | To promote gender inclusivity and to fulfil the main objectives of the project | STREAM IT Website and social media | Public |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Event registration and participation datasets** | Basecamp, Forms, xlsx, | Registration and acceptance of privacy policy of HETFA | Apart from registration there is no foreseen data that will collected | Accomplishing KPIs and quantitative metrics | Gdrive STREAM IT | Only involved STREAM IT Partners and the HETFA |
| **Interview Recording** | Standard recording format | Consent form + asking for approval prior to recording and explaining usage. | By interviewer - only to be used by interviewer | Only for internal use - to facilitate verbatim transcripts of Interviews | STREAM IT Gdrive. Deleted after writing findings have been compiled | STREAM IT partners only |
| **Research Report** | .doc, pdf | Consent form and anonymised | By interviewer. | To compile findings to further project objectives and publish results anonymously | STREAM IT Gdrive Website and social media | Public |
| **Case from interview (text, photos/images)** | .doc , .pdf Photos & images (i.e. .jpg) | By company (via e-mail) ask and integrate feedback prior to approval for usage | Written by interviewer based on interview and information from website. | Inclusion on website and promotion via social media. | STREAM IT Gdrive Approved case on STREAM IT Website | Approved cases will be disseminated, publicly available for awareness raising campaigns. Non approved cases: STREAM IT partners only. |
| **Survey Answers** | Google Forms | Within Survey | Survey distributed to partners | Only for internal use | STREAM IT Gdrive | STREAM IT partners only |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Anonymized Survey Answers** | .doc, .pdf | Information only to be used in an anonymized way | | To compile evidences on impact and participation satisfaction to publish results anonymously | STREAM IT Gdrive Website and social media | Public |
| **Stakeholders List** | Excel, Sheets (*.xlsx) | Internal information only to support initiation of engagement. | | Engagement of stakeholders in the project (ask to engage, inform about GDPR) | STREAM IT Gdrive | STREAM IT partners only. No use outside the intended purpose. |
| **Ext. Presentation** | .ppt, .pdf | Dissemination board | | | STREAM IT Gdrive | STREAM IT partners only. Share only pdf outside. |
| **Publication** | .pdf | Confidentiality obligations and rules of consortium Agreement | | | STREAM IT Gdrive | Public, if approved. |
| **STREAM IT Meeting reports** | .docx video (*.avi, mpeg, mov) | No approval, except if stipulated by consortium agreement. | By reporter of the meeting (and AI tools) | Provide insight to STREAM IT partners on discussed points and actions agreed upon. | STREAM IT GDrive | STREAM IT partners only |

*Table 1: STREAM IT Data Management Table*

## 3. DESCRIPTION OF DATA MANAGEMENT

Research data management within the project adheres strictly to the FAIR principles mandated by the Horizon Europe programme. The dissemination of project data follows a procedure outlined in the Consortium Agreement (CA) and later modified by the General Assembly.

The STREAM IT project is committed to sharing its data openly, subject to ethical and legal considerations, adhering to the FAIR principles. Each partner organisation in the STREAMIT is responsible for complying with the General Data Protection Regulation (GDPR) when collecting data within the WPs they participate in. The project will share data through different channels, including online data repositories, journal publications, and reports. All non-sensitive data will be made available a.o. through the project website, https://streamit-project.eu/.

**Policy on Scientific Publications:**

STREAM IT operates under five guiding principles for managing scientific publications:

- Encouraging early and transparent sharing of research.
- Enforcing active policies for managing research output and data.
- Ensuring reproducibility and originality of research outputs.
- Providing open access to research outputs within the boundaries of Intellectual Property Rights (IPR).
- Involving all pertinent knowledge actors in output processes.

Publication options include:

- Intention for publishing in open-access journals with no Article Processing Charges (APC), open-access journals with "low-cost" APCs, or conference proceedings.
- Opting for gold open access, which involves higher fees for immediate open access.
- Choosing green open access, allowing publication availability after an embargo period.
- Not publishing in open access mode, which is discouraged.

For STREAM IT, these principles are applied as follows:

- All partners are required to provide open access to their scientific publications and data, achieved by granting open access to publications and uploading them to the project website and open access repositories.
- Research outputs will be archived and shared with the EU community through Zenodo or partners' repositories, respecting CA provisions on IP, confidentiality, and publication.
- Methodological aspects in STREAM IT adhere to open science practices while respecting IPR and company interests.
- Partners are encouraged to disseminate outputs on web-based scientific social networks.
- Guidelines outlined ensure compliance with open access obligations, data management rules, and conflict of interest avoidance.
- Default open access and gold open access publication options align with EU GA policies, with pre-prints or accepted manuscripts shared when gold open access isn't feasible.

The project employs a virtual dissemination approach to enhance visibility, adapting materials according to the target audience. This open science strategy aims to:

1. Accelerate innovation by facilitating access to research information.

2. Improve scientific research through increased circulation of ideas.

3. Engage citizens and society, particularly non-profit organisations, by making research openly available, addressing gender representation in STEM.

| ACTIVITIES | DATA TYPE | PROCESSING METHOD | GDPR CONSIDERATIONS |
|---|---|---|---|
| CAPITALIZATION & NEEDS-BASED METHODOLOGY | Knowledge and results of previous and ongoing initiatives | Data will be collected, anonymized, synthesised, and validated. | Anonymize data to protect personal identities. Ensure data minimization. |
| | Research results on institutional barriers | Collect and analyse research results, ensuring anonymity. | Anonymize data to avoid identifying individuals. |
| | STEM education providers' roles | Gather and analyse data from education providers. | Obtain consent if personal data is collected. Anonymize data where possible. |

| | Co-creation Working Lab data | Collect feedback and co-created content. | Ensure all participants provide informed consent. |
|---|---|---|---|
| **GENDER AND DIVERSITY INCLUSIVE TOOLS FOR HANDS-ON ACTIVITIES** | Data from Nurturing Talent Workshop series | Collect participation data, feedback, and outcomes. | Obtain parental consent for minors. Anonymize data for reporting. |
| | Hands-on activities at science centres and museums | Collect participation data and feedback. | Obtain consent and ensure data is anonymized for reporting. |
| | Mentoring Programme data | Collect data on mentor-mentee interactions and outcomes. | Secure consent and ensure confidentiality of personal data. |
| **SUSTAINABLE COLLABORATIVE NETWORKS** | Non-formal teaching and training tools data | Collect usage data and feedback from educators. | Anonymise feedback data. Ensure consent for any personal data collected. |
| | Data from National Inspiration Hubs | Collect participation data and outcomes. | Obtain consent and anonymise data where necessary. |
| | Data from Virtual Makerspace usage | Collect user interaction data and feedback. | Anonymise data and ensure secure storage. |
| **CONTRIBUTING TO WIDER EU OBJECTIVES AND MANIFESTO** | Impact assessment data | Collect and analyse data on project impact. | Anonymise data to protect individual identities. |
| | Policy recommendation data | Gather and synthesise data for policy recommendations. | Ensure any personal data used is anonymized. |
| **AWARENESS RAISING AND ENGAGEMENT OF YOUNG PEOPLE** | Communication and Dissemination Strategy data | Collect data on communication strategy effectiveness. | Anonymize data and ensure secure handling. |
| | Awareness raising campaigns data | Collect participation and engagement data. | Obtain consent where personal data is collected. |
| | Engagement of target groups data | Collect feedback and engagement metrics. | Anonymize data and ensure consent is obtained. |
| | Data on Project Ambassadors activities | Collect data on ambassador activities and impact. | Ensure ambassadors' data is processed |

| | | | with consent and anonymized. |
|---|---|---|---|

*Table 2: Work Package data summary*

## 4. FAIR DATA MANAGEMENT

### 4.1 MAKING DATA FINDABLE, INCLUDING PROVISIONS FOR METADATA

For identification purposes, the STREAM IT consortium will follow specific naming conventions for all data sets, including date, project name, WP, relevant Task, Deliverable or Milestone and version number. To ensure documents are easily findable within repositories, metadata will be attached to each document. This includes mandatory metadata, such as:

- European Union
- Horizon Europe
- HE
- STREAM IT
- GA No. 101131843

Specific metadata will be defined according to the research data's content. Examples include:

- Educators
- STEAM
- STEM
- Schools
- Higher Education
- Capacity building
- Webinars
- Policy dialogue

All data and metadata will be stored in English to enhance their findability within the scientific community. For identification purposes, repositories offer persistent and unique identifiers like Digital Object Identifiers (DOI). ZENODO, for instance, supports DOI versioning for document updates.

### 4.2 MAKING DATA OPENLY ACCESSIBLE

The results of STREAM IT project activities will be summarised in deliverables and reports submitted to the European Union. Public deliverables, communication, and dissemination materials will be openly accessible on the STREAM IT project website. Personal and company data will remain confidential in compliance with GDPR.

STREAM IT data that can be shared with the public will be identifiable and locatable using standard identification mechanisms, such as DOIs and search keywords. Each partner is

responsible for making their data from the STREAM IT project open according to HE FAIR guidelines. Data will be stored in institutional repositories, the STREAM IT repository, or in ZENODO. Data uploaded to ZENODO will be linked to OpenAIRE and the EC portal, ensuring public accessibility.

In addition to these repositories, findings and research can be shared on social networks like LinkedIn and the project website at https://streamit-project.eu/.

The procedure will be as follows: once project results are published, PDF copies and any complementary data will be uploaded to the selected repository. They will be announced on the website with links to the publication location and disseminated via newsletters, conferences, LinkedIn, and Twitter to reach a broad audience.

### 4.3 Making Data Interoperable

Data produced in the STREAM IT project will be interoperable, allowing data exchange and re-use between researchers, institutions, organisations, and countries. Public data will adhere to standards for formats, compatible with available (open) software applications, facilitating re-combinations with different datasets from various origins. No specific software tools are required to access STREAM IT data; common software such as the GSuite, Microsoft Word, Excel, Adobe Acrobat Reader, or Open Office will suffice.

Public press releases and social media news on LinkedIn and Twitter will use common language for public understanding.

### 4.4 Increase Data Re-use

STREAM IT data can be shared and re-used, except for personal data, which will be treated according to GDPR, and business-critical data. Data in public repositories may be used by third parties for research purposes, avoiding duplication of efforts and providing a basis for future investigations. The generated data can be re-used in similar configurations.

Commercial use of project-generated data is restricted to consortium members only.

## 5. Allocation of Resources

Data preservation responsibility lies with the partner who produced the data. The coordinator gathers and reports documents to the EU. Each partner is responsible for making its data and results open and uploading them to repositories. These costs are eligible for reimbursement during the project duration. The coordinator HETFA is responsible for updating the DMP together with Steinbeis. Costs for documentation preparation and uploading should be included in WP1 "Project management" and are eligible for reimbursement under EU rules. Initially, free repositories like institutional ones and ZENODO will be used.

## 6. DATA SECURITY

Project data will be deposited within the described repositories. Data files will be managed, processed, and stored securely (e.g., password-protected systems, firewalls, surge protection, and virus protection). Consortium members have backup systems to preserve their information.

## 7. ETHICAL ASPECTS

The STREAM IT project will adhere to strict ethical standards for data management. The project will ensure that all data is collected, stored, and shared in compliance with the EU General Data Protection Regulation (GDPR):

1) Personal and company data will be kept confidential as required by GDPR.
2) Data collected will be handled according to partner privacy policies, all committed to GDPR compliance. Participants will always be provided with a link to the related privacy policy when personal data is collected.
3) Ethical consideration for Survey: Answering to the survey is voluntary and all data will be collected anonymously by setting the survey tools to not collect IP addresses. Email addresses will be only gathered from those who are interested in learning about the results and/or volunteer for interviews and/or if they agree to share more information as training providers. These data will be handled in accordance with GDPR. None of the personal data will ever be passed to third parties.
4) Ethical consideration for Interview: Participation in the interview is voluntary. The interviewees must sign a consent form before the interview. Interviewees will remain anonymous throughout data reported. None of the personal data will ever be passed to third parties. Templates for informed consent forms and information sheets covering voluntary participation and data protection issues are provided to any participant in STREAM IT activities by HETFA.
5) Ethical consideration for Webinar: Webinars will be established online, dedicated to educators/educator providers. GDPR rules and concerns will be applied to the organisation of the webinar events.

## 8. OTHER

STREAM IT does not use any other national/funder/sectorial/departmental procedures for data management. Each institution has implemented procedures to guarantee data preservation and curation, aligning well with EU guideline

# STREAM IT