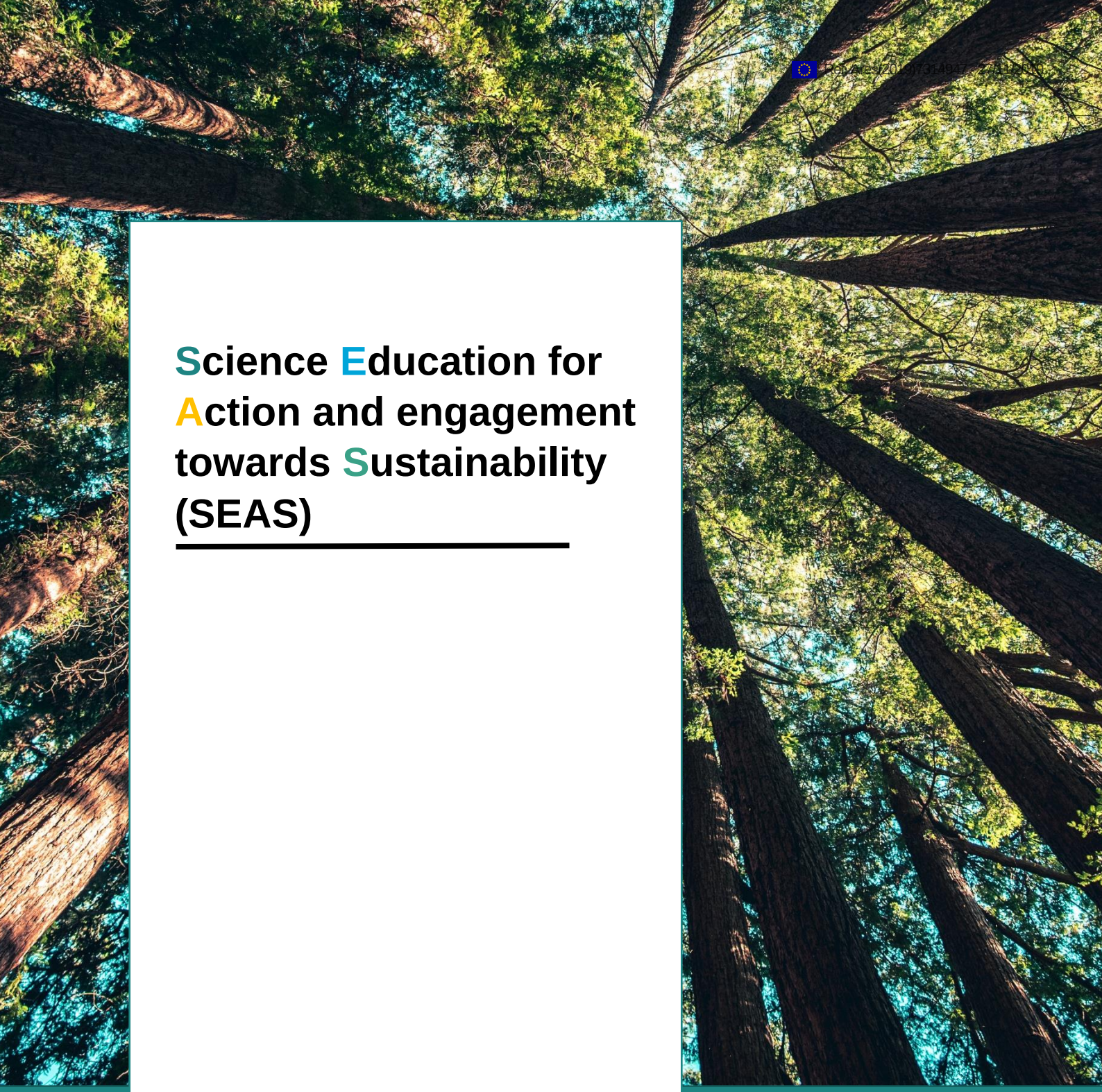


Science Education for Action and engagement towards Sustainability (SEAS)

D7.1.

Report on Data Management Plan

Project no.	824522
Project acronym:	SEAS
Project title:	Science education for action and engagement towards sustainability
Funded under:	SwafS-01-2018-2019-2020 - Open schooling and collaboration on science education
Document identifier:	D 7.1. Report on Data Management Plan
Version:	1.0
Submission date:	28.11.2019
Work package:	WP 7 - Management
Dissemination level:	Public
Lead beneficiary:	UiO

Author list	
Deliverable leader [name] [e-mail]	*Erik Knain, erik.knain@ils.uio.no*
Name of author(s)	*Name of organisation*
Erik Knain	University of Oslo
Alfredo Jornet	University of Oslo
Torgeir Christiansen	University of Oslo

Acknowledgements

Thanks to Jelena Radisic for comments regarding questionnaire data and consent form, and to Susanne Kubisch for important comments in the internal quality assurance process.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824522.

Contents

1 Summary 3

2 Document history and co-authorship 4

3 Introduction 4

4 Data Summary 5

4.1 The Role of Data in SEAS 5

4.2 Sources and Flow of Data in the SEAS WP Structure 7

4.3 Roles Connected to Data Flow in SEAS 8

4.4 Data Types, Formats and Purpose 10

4.5 Use of Existing Data in SEAS 13

4.6 Data "Utility" 13

5 FAIR data 14

5.1 Making data findable, including provisions for metadata 14

5.2 Folder structure in SEAS TSD 14

5.3 File name structure in SEAS 16

5.4 Data log file(s) 20

5.5 Log relating codes defined within each local network 21

5.6 Log file connecting data files and participants' consent. 21

5.7 Making data openly accessible and increase re-use 21

5.8 Making data interoperable 22

6 Allocation of resources 22

7 Data security 23

8 Ethical aspects 24

8.1 Ethical requirements (Post-grant requirements) – WP 1 24

Appendix 1: Information and consent generic document 25

1 Summary

The collection and analysis of data are essential to the development of research-based practices, insights, and models of scientific literacy for sustainable citizenship and open schooling, which are at the heart of SEAS. The Data Management Plan (DMP) is a report, and a guideline, about the use and storage of the different types of data that will be generated throughout the SEAS project. The report describes the data life cycle, which includes the first contact with the participants in the settings where data are to be collected, the informed consent requirements, the collection and types of data, their processing, storage and sharing, and procedures for sharing and disseminate the outputs from the analyses of the data.

There are two main sources of data in SEAS. The first one is the local assessment procedure, which consists of a set of generic concepts, methods and approaches that are to be implemented through qualitative and quantitative methodologies in the local networks. The second main source of data in SEAS is the global assessment instrument, a questionnaire aiming to assess by quantitative measures key areas of competence that SEAS aims to have an impact upon, including scientific literacy, agency, and interest in science.

The SEAS data and resulting documentation, in turn, become part of the projects' FAIR strategy (making data Findable, Accessible, Interoperable and Reusable). To achieve FAIR data within SEAS while strictly following ethical rules in accordance with GDPR, SEAS utilizes the TSD (Services for Sensitive Data), a service for secure digital storage of data located at and managed by the University of Oslo. In the document we describe a set of measures that will be deployed to ensure FAIR access within SEAS, including:

a) folder structure
b) naming of files
c) a structured log file overseeing the entire database
d) a log relating codes defined within each local network, and
e) a log file connecting data files and participants' consent

With respect to measures to ensure FAIR access to a wider community of users outside the SEAS consortium, long term depository is planned to be serviced by the Norwegian Centre for Research Data (NSD) who is a certified infrastructure for archiving and sharing data. A later version of the DMP will describe a possible core for metadata for archived data.

SEAS publications will adhere to green and gold open access publicing, preferably gold open access.

The DMP is based on the H2020 template, structurally, and addresses the questions set out in that template. DMP is an evolving document that will be updated after each iteration in SEAS (AUG20, JUN21, DES21, and JUN22).

2 Document history and co-authorship

The SEAS Data Management Plan (DMP) is an evolving document that will develop throughout the SEAS project lifetime. The first version is developed based on (a) Section 5.1 of the SEAS' project description, and on (b) work meetings between the three authors at the University of Oslo. This deliverable was send to the internal quality control team on November 18.

Version	Date	Notes
1.0	28.11.2019	Initial version

3 Introduction

The collection and analysis of data are essential to the development of research-based practices, insights, and models of scientific literacy for sustainable citizenship and open schooling, which are at the heart of SEAS. To reach these main objectives, the SEAS project includes the establishment, support and coordination, as well as the evaluation of six local open-schooling networks in European countries. These networks will be diverse, thus allowing to develop general principles, values and goals according to the local foci and needs, emerging from local negotiation and democratizing inquiry. To ensure both flexibility to accommodate local needs and idiosyncrasies, while at the same time generating general knowledge and models, the project develops a shared pool of ideas, concepts, values, tools and methods, including data collection methods, that are shared across the local networks (see deliverables *D2.1 Plan and milestones concerning challenge implementation for each local network*, *D2.2 A document with the shared pool of SEAS' analytical concepts and methods*, and *D2.3 Guideline documentation for data collection and local assessment procedure*). During the project period, these initial framing documents are to be developed into more generic, transferrable knowledge and insights, through the SEAS' iterative design-based research approach, and its methods for local and global assessment. A range of data is needed in order to generate these insights.

The DMP is a narrative, and a guideline, about the use and storage of different data sets that shall be generated throughout the SEAS project. The report describes the data life cycle, which includes the first contact with the participants in the settings where data are to be collected, the informed consent requirements, the collection and data types, their processing, storage and sharing, as well as procedures for disseminating the outputs from transformations and analyses of these data.

Assessment methods, selected data and results from analysis at local network level and global project level will be shared through deliverables and publications, both during the project period (including two annual reports on local assessment and best practices, *D3.1.* and *D3.3.*), and as final outcomes of SEAS (including *D5.2. Report on models for the establishment and implementation of open schooling partnerships*, *D5.3. Report on models on transformational engagement scientific literacies and motivation*, and *D5.4. Report on models of teaching and learning scientific literacy*). This data and resulting documentation, in turn, become part of the projects' FAIR strategy (making data Findable, Accessible, Interoperable and Reusable) (see section 5 for a detailed description of SEAS' FAIR strategy). However, the FAIR principles need to be observed with respect to requirements deriving from research ethics and GDPR regulations. The purpose of this document is to operationalize these diverse concerns on accessibility, privacy, and usability, in a set of principles and procedures to be observed within the SEAS Consortium, and for potential users and stakeholders outside the Consortium.

As part of the European Commission's Open Research Data Pilot (ORD pilot), SEAS is required to develop and update a DMP, storing data in a data repository that third parties shall be able to access and providing them with the tools needed to freely access, mine, exploit, reproduce and disseminate data, as well as to validate the project's research findings.[1] Research ethics will however enforce limitations to the ORD principles, and, as the H2020 online manual notes, "Participating in the ORD Pilot does not necessarily mean opening up all your research data." Rather, the ORD pilot follows the principle, ***as open as possible, as closed as necessary,*** and encourages sound data management. The DMP also complies with the sections regulating ownership, transfer and dissemination of results in the SEAS Consortium Agreement.

The DMP is based on the H2020 template, structurally, and addresses the questions set out in that template. DMP will be updated after each iteration in SEAS (AUG20, JUN21, DES21, and JUN22).

4 Data Summary

4.1 The Role of Data in SEAS

The main objective of the Science Education for Action and Engagement towards Sustainability (SEAS) project is to establish, coordinate and evaluate collaboration among *six open schooling*

[1] https://www.openaire.eu/what-is-the-open-research-data-pilot.

networks led by universities and science centres with partners in Austria, Belgium, Estonia, Italy, Norway, Sweden, and the UK. As part of these networks, students in formal schools are given the opportunity to engage in addressing real-life, complex sustainability challenges that are identified and dealt with together with participants and stakeholders in the local community. Ultimately, SEAS will develop models for open-schooling implementation that will support schools and out-of-school institutions to collaborate in the promotion of 21st century scientific literacy and skills for responsible citizenship among students and citizens in the larger community.

To achieve these general goals, the specific objectives of SEAS are operationalized as follows:

1. To create six open schooling networks that support engagement with science for action in real-life complex sustainability challenges, that are relevant for partners of formal, informal and non-formal institutional settings
2. To develop digital resources at the local community level and project level that support in-person meetings, collaboration, and exploration
3. To broaden science education to increase its relevance for students and to enhance greater interest in science, and thereby also increase the chances that students choose future careers in science
4. To develop an assessment framework across formal, informal and non-formal science education with an aim of providing quality assurance in future open schooling initiatives, and to assess the effects of such initiatives
5. To implement and assess tools and resources that are differentiated according to institution, purpose, and users that support the development of scientific literacy
6. To differentiate the tools, strategies and design principles taking into consideration gender differences and students' backgrounds, as well as geographical variation
7. To develop a training program for teachers and teacher educators
8. To provide a white paper on policy recommendations for open schooling in a European context
9. To develop a platform for dissemination that opens for tools to be configured and adapted according to users' needs

To complete its objectives, SEAS develops, explores, and adjusts according to local needs, a set of concepts, digital tools, and methods for enabling and assessing open-schooling science learning collaborations across formal, non-formal and informal settings.

Challenges and achievements in the local networks will be documented and assessed both with the purpose of informing the project's implementation and for generating transferrable knowledge, models, and training materials that will inform and promote future open schooling implementations. These funds of knowledge, tools, and models will help overcoming documented challenges related to knowledge transitions across contexts, building new institutional means and processes for supporting science learning that will shape scientific knowledge into more usable and transferable forms in and for contemporary society.

4.2 Sources and Flow of Data in the SEAS WP Structure

There are two main sources of data in SEAS. The first one is the local assessment procedure, which consists of a set of generic concepts, methods and approaches that are to be implemented through a variety of qualitative and quantitative methodologies in the local networks. A generic description of local assessment is to be found in *Deliverable 2.3 "Guideline documentation for data collection and local assessment procedure"*, which will have the status of confidential for data integrity purposes until the data collection period finishes. The aim of the local assessment procedure, implemented by WP3, is to identify best-practice examples and challenges, in order to further inform the development of open schooling practices at the local level. Further aim is to feed into the generation of more general insights at the global synthesis level (WP5).

The second main source of data in SEAS is the global assessment instrument, a questionnaire aiming to assess by quantitative methods key areas of competence that SEAS aims to have an impact upon, including scientific literacy, agency, and interest in science. These areas of competence, as well as other practical aspects of the global assessment instrument are described in detail in deliverable *D5.1. A global assessment instrument document*, which will be kept confidential until end of data collection to prevent dissemination of methodology interfering with the instrument's validity.

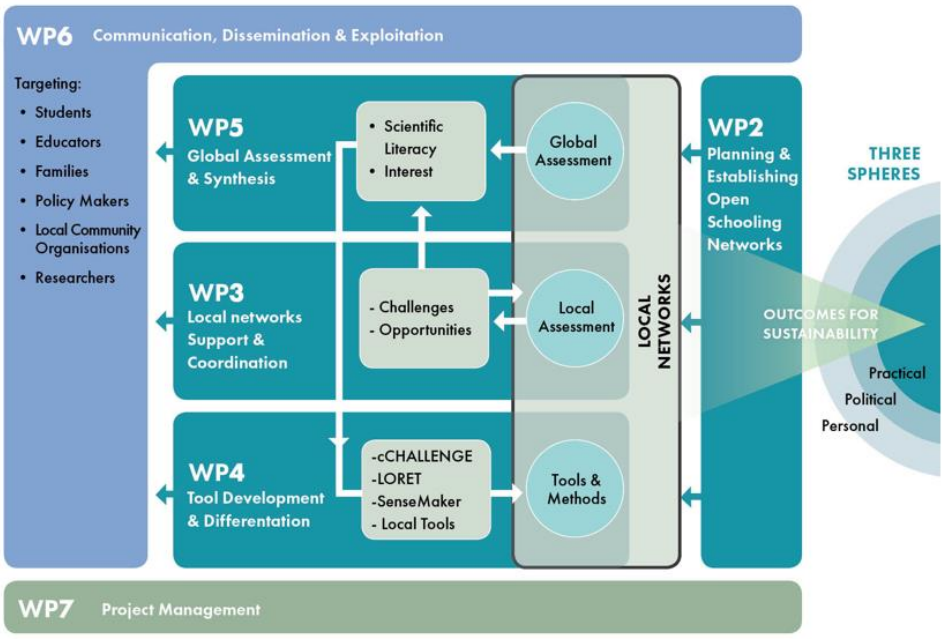


Figure 1: The SEAS work package structure and data feedback flow across WPs.

Data and analyses generated through both global assessment instrument (WP5) and the local assessment procedures (WP3), are then further analysed and developed for synthesis in WP5, where a more generic analysis including global and local data collected across all local networks, is synthesized with the purpose of generating transferrable knowledge. This transferrable knowledge is then brought back to the implementation of open schooling within the SEAS project through WP4, where existing tools are updated, or not-yet-existing tools generated, to address ongoing needs

and knowledge. Transferrable knowledge is also taken up at the local network level via reconfiguration of concepts and methods generated and sustained in WP3 (see Figure 1 for an illustration of the work packages structure and flow of data/information).

The flow of data and knowledge described above means that data will need to be made accessible across partners within the consortium during the project period, following a FAIR strategy, as described above, but also attending to ethics (e.g., consent agreement) and to overall GDPR guidelines. To achieve FAIR data within SEAS while strictly following ethical rules in accordance with GDPR, SEAS utilizes the TSD (Services for Sensitive Data), a service for secure digital storage of data located at and managed by the University of Oslo. Throughout the DMP, we return to TSD frequently. But as a short introduction, TSD provides a virtual space containing some pre-specified functions and items for the secure storage and processing of data. These pre-specified functions and items are further enhanced with SEAS-specific folders and functions that are tailored to SEAS' goals and methods.

4.3 Roles Connected to Data Flow in SEAS

In order to preserve the FAIR and GDPR goals and requirements described above, a set of procedures for moving data from data collection to storage and sharing are clarified in this document. These routines imply the need to appoint two main types of roles: Local data manager, and overall data flow manager. The overall data flow manager is the leader of WP3, which is in charge of collecting and archiving the SEAS data base. Prior to the start of data collection, each local network leader must assign the role of local data manager to a person of the team, who will be responsible for ensuring that the data management procedures and routines described here are followed.

The routines described below include two key elements:

- Data collection tasks. A delimited data collection task is established at the local level by following guidelines in *D2.3 Guideline documentation for data collection and local assessment procedure*, and in deliverable *D5.1 A global assessment instrument document*. Each data collection task is delimited in terms of setting, time, and data type and purpose.

- SEAS local data package. Each collection task leads to one or multiple *data packages*. Data packages are defined locally in terms of convenience with respect to how a set of diverse data types or formats may be internally connected with respect to setting, time, and/or purpose. For example, a same data collection task may be run with 3 different schools, in which case, it may be convenient, for practical and security purposes, to collect the diverse data together into a package that is delivered (uploaded) at once. Data packages include not only (a) data, but also (b) relevant data log, (c) a log with the relation of codes given locally, and (d) consent forms related to the data being uploaded.

The routines to ensure a secure data flow can be listed as follows:

1. Data collection is organized following guidelines described in deliverable *D2.3 Guideline documentation for data collection and local assessment procedure*, and in deliverable *D5.1 A global assessment instrument document*. These two documents establish local and global assessment data collection requirements, which specify given *data collection tasks*. Each data collection task is delimited in terms of setting, time, and data type and purpose. Each collection task, in turn, leads to one or multiple *data packages*. Data packages are temporarily stored following requirements specified by each local approved ethics and consent forms.

2. Local data manager collects data in data packages and ensures that all relevant data and documentation is appropriately named and included. This includes ensuring that the relevant consent forms, and a log file connecting files and consent forms is included (see details in section 5.2 on FAIR data access and sharing). This also includes ensuring that a written Data recording log of incidents during the data collection in the field is made and then checked for incidents (for instance that someone without consent is recorded). Furthermore, it includes ensuring that the files are properly named following the instructions provided in section 5.3 on file name structure. Once all files are ready, the local data manager creates a new folder with the work package name under the SEAS TSD folder "Inbox" (see SEAS TSD folder structure in section 5.2.). Each country has a dedicated "Inbox" folder where they (and only they) can upload files. The inbox folder is the only folder where local partners have rights to modify files. Corrective action in case incidents are registered on the data log of incidents during data collection must be taken, prior to the uploading to the Inbox folder or, at the latest, prior to informing the overall data flow manager that the data package is uploaded (stage 3).

3. Once a new folder is created inside the Inbox folder, and a data package is uploaded, the local data manager informs the overall data flow manager of the upload via e-mail and/or a message using the SEAS Teams platform.

4. Overall data flow manager is responsible for checking that the data package delivered by local data manager is complete and adequately uploaded. Once quality is checked, the data files in the data package are moved over to the permanent "Data" folder within SEAS TSD, while the log files and consent forms are added to the data log folder and the consent forms folder, respectively. Once there, all partners have read access, while only the overall data flow manager has access to write and modify files. Once the data are stored permanently, the Inbox folder must be emptied.

5. A dedicated export folder in TSD allows data to be shared with analysists outside the local network domain in a separate domain in TSD. Only the overall data manager in SEAS (the Coordinator) or the overall data flow manager in SEAS have permission to conduct the file transfer from the export folder to the shared work space.

4.4 Data Types, Formats and Purpose

As specified, there are two main sources of data in SEAS, the local assessment procedure, and the global assessment instrument. However, both these sources shall include a variety of data. Generally, the data collected in SEAS will include text, video, audio, still photos, SenseMaker entries (web solution), cCHALLENGE entries (web solution), log files and questionnaire data. Applicable open or commercial standards for video, audio and still photos will be used for formatting. Web solutions will be stored as text (XML, HTML, JPG, PDF) as far as possible. Log files will be stored as text or CSV files, or similar.

When files are compressed to enclose raw data into a usable (standard) format, this will be done without loss of data where possible. For example, when some photo cameras use in-house encoding that may not be accessible to other partners/users in the consortium, the image files shall be converted into an open or industry standard format (like JPEG, TIFF, or PNG). However, different methods of compression/conversion may exclude relevant data from the raw files. Judging which aspects of the data may be essential to preserve is important. The purpose of this section is to assist SEAS members to consider the possible purposes and uses of a given file before converting it for sharing and storage.

In order to provide guidance with respect to possible format and processing choices, Table 1, presents a list of data types, purposes, and associated format/process considerations that are expected to be relevant in the SEAS data management. The typology is based on the context of data collection known to be going to take place. However, additional types of data may be identified based on new emerging contexts, or when relevant for the implementation of the project's local and global assessments strategies, taking due consideration to consent. In line with SEAS' iterative design process, where practice is refined and further developed throughout experimental iterations, it is possible that additional data types are specified after this version of the DMP is published.

Data Type	Description and Purpose	Format/process considerations
Interviews	Interviews are generally used for better understanding participants' opinions, prior experiences. Further, they are used for events, possibly also events documented through other data collection forms that now can be deepened through interview. Interviews may be conducted with teachers,	Audio recording may be enough for one-to-one interviews, while video recording may be more appropriate in cases where there are multiple participants or where the context (setting) of the conversation is important and may be difficult to retrieve from audio only. Standard audio formats include (but are not limited to) WAV and mp3. Standard

	students, family members, or other stakeholders; individually or in groups.	video formats include (but are not limited to) AVI, MOV and MP4.
Meetings and Workshops	Local network meetings and ChangeLab Workshops shall be documented through ethnographic methods of data collection, and preferably using video. Notetaking, observation protocols, pictures, and audio recordings may also be relevant. The purpose of documenting these meetings and networks is to study processes of collaboration in open schooling networks, both within and across institutional contexts.	Video recording with multiple video angles and audio sources is preferred. Standard audio formats include (but are not limited to) WAV and mp3. Standard video formats include (but are not limited to) AVI, MOV and MP4. Field notes should be digitalized into readable text file, such as txt, rtf, or docx. PDF may also be adequate if automated text recognition is run or if text image is good enough for such a recognition to be run. Images of settings, student and teacher tools and products, etc, are to be uploaded as JPEG, TIFF, or PNG
In-school activities	Activities carried in the school (the classroom, or other areas of the school setting) as part of the SEAS sustainability challenges shall be documented through ethnographic methods of data collection, and preferably using video. Notetaking, observation protocols, pictures, and audio recordings may also be relevant. The aim of documenting in-school activities relates to an interest in understanding both the challenges and opportunities for school teaching and learning that emerge as part of open schooling projects, and to document emerging forms of	Video recording with multiple video angles and audio sources is preferred. Standard audio formats include (but are not limited to) WAV and mp3. Standard video formats include (but are not limited to) AVI, MOV and MP4. Field notes should be digitalized into readable text file, such as txt, rtf, or docx. PDF may also be adequate if automated text recognition is run or if text image is good enough for such a recognition to be run. Images of settings, student and teacher tools and products, etc, are to be uploaded as JPEG, TIFF, or PNG

	scientific literacy along with possible narratives of change.	
Out-of-school activities	The same description as in in-school activities holds for out-of-school activities, where the shift is with regards to setting. Documenting activities taking place outside of the school, has the additional purpose to document challenges and opportunities for transformative science learning that emerge at the boundary of institutional practices.	Video recording with multiple video angles and audio sources is preferred. Standard audio formats include (but are not limited to) WAV and mp3. Standard video formats include (but are not limited to) AVI, MOV and MP4. Field notes should be digitalized into searchable text file, such as txt, rtf, or docx. PDF may also be adequate if automated text recognition is run or if text image is good enough for such a recognition to be run. Images of settings, student and teacher tools and products, etc, are to be uploaded as JPEG, TIFF, or PNG
Non-School Learning	To fully understand the impact and significance of open-schooling initiative in the learning lives of the participants, it is necessary also to document learning trajectories taking place outside of schooling time. This includes time at home, with friends, and any sort of after-school activity. Methods of data collection here include personal diaries and self-reports that participants can implement on their own and deliver as research material.	Participant generated diaries and other forms of self-reporting, documenting the participants' learning lives outside of schooling time are to be digitalized via scanner or other forms of digital copy. They should be digitalized into searchable text file, such as txt, rtf, or docx. PDF may also be adequate if automated text recognition is run or if text image is good enough for such a recognition to be run. Images of settings, student and teacher tools and products, etc, are to be uploaded as JPEG, TIFF, or PNG
Digital log activities	An important part of learners' life as students takes place in the digital domain, with some digital tools being specifically tailored and used as part of SEAS. To document the type of discourses, routines, and	Different software applications offer different opportunities and formats for logging. Digital logs must be planned in accordance with instructions in deliverable *D4.1 guideline manual describing use of each of the SEAS digital tools*

	emerging narratives of change, participants' activities in the digital domain may be logged for selected periods/samples/tools.	
Questionnaire items, quantitative	Individuals' scores on diverse measures concerning key areas of competence of interest to SEAS will be collected through the global assessment instrument.	The questionaire will be administered to respondents through online platform LImeSurvey. Data can be exported to be readable in R, Stata, SPSS and as excel sheets. A more detailed description is to be found in *D5.1 A global assessment document*.
Questionnaire items, qualitative	In addition to quantitative items, the global assessment instrument includes open questions.	Students type answers in a single text cell which can be exported to be readable in excel or SPSS.

Table 1: Data types and their purposes in SEAS

4.5 Use of Existing Data in SEAS

In the event that existing data will be used in SEAS, such use will comply with the regulations in the Consortium Agreement. Any use of existing data must conform to 1. the consent and ethics of the data in question, and 2. that the ethics in the generation of data in question aligns with SEAS' ethics; 2. Existing data must be stored separately from the SEAS collected data and be fully traceable according to ethics and FAIR principles.

4.6 Data "Utility"

As pointed out above, data is collected, assessed, and analysed in SEAS with a dual purpose in terms of users. One purpose is to inform the project's implementation by supporting and developing the local networks in SEAS. This means that the data collected will be directly useful for all participants in SEAS. A second purpose is to generate transferrable knowledge, models, and training materials that will inform and promote future open schooling implementations. Thus the data collected and analysed by the SEAS consortium will be useful for practitioners and scholars interested in the processes involved in open schooling, science education for citizenship and sustainability education as operationalized in SEAS.

5 FAIR data

5.1 Making data findable, including provisions for metadata

SEAS considers FAIR principles at two levels, that is, in terms of two types of user: addressing users within SEAS, which include partners within the consortium, and addressing users outside of the consortium, reaching out towards a wider community of practitioners and scholars. The main focus of this version of the DMP (1.0) is the former, as it is the most pressing requisite for achieving the most immediate goals of the project, namely, documenting practices in and for the open schooling local networks. Below, we describe a set of measures that will be deployed to ensure FAIR access within SEAS, including:

f) folder structure, (section 5.2)
g) naming of files, (section 5.3)
h) a structured log file overseeing the entire database,
i) a log relating codes defined within each local network, and
j) a log file connecting data files and participants' consent.

These measures are described in detail in the following sections.

With respect to measures to ensure FAIR access to a wider community of users outside the SEAS consortium, these will include anonymized transcriptions of video data, video-and sound data for which consent has been provided for audio-visual materials that can be used in presentations and articles in open access media, implied by the widest option in consent form. A later version DMP will describe a possible core for metadata. The Dublin Core[2] is one possibility. The 2012 Version 1.1 of the Dublin Core includes the following elements: Contributor (entity responsible for making contributions to the resource), Coverage (spatial or temporal), Creator (entity responsible for creating the resource), Date, Description, Format, Identifier, Language, Publisher, Relation (related resource), Rights, Source (related resource from which the described resource is derived), Subject (topic of resource), Title, and Type. As SEAS local assessment data will be highly contextual, it will likely be necessary to include additional elements to the choice of a metadata standard such as the Dublin Core in order to make the data become findable by data base searches. Possible option is to utilize the Description element in the Dublin Core to specify contextual information or the nature of the interaction (for instance whole class/group, type of interaction and subject matter theme).

5.2 Folder structure in SEAS TSD

As a means to store and make the SEAS data accessible in ways that satisfy both the GDPR requirements and other ethics requirements related to research data privacy and security, all data

[2] https://www.dublincore.org/specifications/dublin-core/dces/

collected as part of the SEAS project will be stored, accessed and processed within the TSD virtual space (a detailed description of how to access and use TSD will be found in SEAS – General).

The SEAS TSD folder structure reflects different functions as part of the data management flow. The folder structure is nearly flat (almost no folder hierarchies), and includes the following items (folders):

- Inbox folder
- Data folder
- Data log folder
- Consent Forms
- Processed Data
- Export folder

Below, we describe each of these items, and their role in the overall data flow.

Inbox folder

The inbox folder is the folder in which data can be uploaded by the different local networks. No files can be uploaded to any of the other folders directly; only the inbox folder is the incoming data folder. Local data managers are responsible for uploading data packages to this folder. All data collected according to the principles and routines described in the DMP must be uploaded, i.e. the data sets must be complete. Each country has access to upload to an inbox folder dedicated to that country, such that there are 6 folders under Inbox folder named as [country inbox], e.g., Austria Inbox, etc.

Data folder

Data files received in the inbox folder are moved over to the permanent data folder, after quality check. To ensure security and accountability, once stored here, these files cannot be edited or deleted by anyone else than the overall data flow manager, and the project data manager. The structure within this folder is flat, with the file name structure being the key to organize the files within this folder (see section 5.3 below on file name structure in SEAS).

Data log folder

The data log folder contains documentation that provides an overview of the contents, and which will provide keys for users to search for and find data relevant to their respective research purposes. Files that provide an overview of the data will contain not only core information such as which files contain which type of data in which types of settings, etc., but will also provide detailed information concerning contents and processes recorded relevant to the SEAS goals. The specific number and structure of these files, which can be anticipated to be spreadsheets such as Microsoft Office Excel, will be developed throughout the project life and monitored in collaboration by both local data

managers and overall data flow manager. One of these files will be a log file specifically dedicated to provide a relation of files and respective consent forms, which will make possible for researchers to trace back data for specific participants or groups of participants if there were changes or specific requests associated to specific consent.

Consent forms folder

The consent forms folder contains all the consent forms collected that are the sole basis that grants the researchers rights to manage the data collected. Each country will have a dedicated folder and read access to that folder, while only overall data flow manager and project data manager have access to all of them. Consent forms are stored separately from the data files, and the key to trace which data belongs to which consent is given in a specific log file stored in the data log folder.

Processed data

Users will be able to process data within the SEAS TSD virtual space by means of dedicated software accessible within that space. The outcomes of these products will be stored in the processed data folder, where all partners have access and admin rights. Processed data should always be anonymized, with a log file describing the specific data types, settings, etc., at the source of the stored processed results.

Export folder

A dedicated export folder in TSD allows data to be shared with analysists outside the local network domain. The local data manager first transfers data that is to be shared into the export folder. These data are then moved from the export folder to a dedicated shared folder (still within TSD) where they can be accessed by SEAS team members who has access to the shared folder. Only the overall data manager in SEAS (the Coordinator) or the overall data flow manager in SEAS have permission to conduct the file transfer from the export folder to the shared work space.

5.3 File name structure in SEAS

A crucial measure for making data findable and for including metadata information involves consistently following a given file naming convention. All data collected as part of the SEAS project will be securely stored in TSD (Services for Sensitive Data), a service for secure digital storage of data located at and managed by the University of Oslo. To ensure the maintenance of data integrity, and the cataloguing of data for accessibility within the TSD space, each file stored in the SEAS project's database is given a unique name according to the following formula:

Network_Date_DataType_School#_Device_file#.Format

In the following, we describe each of the elements/codes in the name convention:

Network refers to the specific local network as part of which the data being stored has been collected. Each country is in principle expected to have one SEAS local network, but, not to exclude

the possibility that a country generates more than one network, the code provides for the possibility of having multiple networks within a country. The network code thus is given by two letters indicating the country of the network, and a number indicating which network within that country the data refers to. The code follows international conventions for the abbreviation of country names, such that, for example, to store a file collected or generated as part of the local network in Norway, the first element of the name would be NO1. If there was a second local network (i.e., a network involving different or partially different partners at the local level with its own sustainable challenges and/or work settings/structure), the code used would be NO2. The network element, therefore, follows the formula [two letters international country abbreviation + number of network within country].

International abbreviations for the countries involved in SEAS are listed below:

Country	Code
Austria	AT
Belgium	BE
Estonia	EE
Italy	IT
Norway	NO
Sweden	SE

Table 2: Relation of country codes

Date

We follow international standard for writing date as year-month-day, but as a single digit (without dash), that is, YYYYMMDD.

If the data captured in the file to be stored would have been collected the third of January of 2020, in the Norwegian Local network 1, then the beginning of the file should read as follows: NO1_20200103

Data Type

As explained in section 4.4, data types in the SEAS project are categorized in terms of the nature of the setting/activity being documented. Table 1 lists the types of data that are expected to be addressed throughout the project's data collection. As SEAS is an innovation project, however, and

the very assessment procedures underlying its development are designed to inform back to the project's research design, we anticipate that it is possible, and perhaps also likely, that new types of settings/activities will be identified. Data types, therefore, will certainly include, but will not necessarily be limited to the ones described in this document.

To make type of data immediately visible, we use word shortenings as code. Table 3, below, presents a relation of codes for types of data.

Data Type	Code
Interviews	int
Meetings and workshops	wks
In-school activities	insc
Out-of-school activities	outsc
Non-school activities	nonsc
Qustionnaire	qs
Digital log	log

Table 3: Relation of data type codes

Thus, if the data to be stored is an interview, recorded the third of January of 2020 in the Norwegian local network 1, the first elements of the file name would read as follows: N01_20200103_int.

While the data types related here are quite generic, a more detailed description of the specific activities being documented in each file will be included in the structured log file overseeing the entire database, as described in section 5.2.

School

School# specifies the number identifying each school in each local network. Schools are identified with a single two-number digit (e.g., 01, 02, 03, …, 10, 11…), with the lowest number being 01. Each local network may have multiple schools. Each local network assigns a two-number digit to each of the schools participating in that network, and uses that number consistently throughout the project period. The relation between school digits and school names within each network is provided in the local networks' codes log file (see description below, section 5.4).

Note that, even when the type of data collected may concern activities unrelated to the school (e.g., data types "nonsc," "int," or "wks") there may be cases in which the data collected concern a particular student or group of students belonging to one school. In that case, the code for that school is used even when the activities documented are not part of the school's activities.

However, there may be cases in which (i) no specific school, or even (ii) multiple schools are involved. Examples of the former include interviews with partners or stakeholders not connected to any given school. Examples of the latter include meetings or workshops in which multiple schools are participating. In those cases, the code 00 (zero zero) is used.

If, say, the data to be stored is an interview, recorded the third of January of 2020 in the Norwegian local network 1, with one of the teachers of school 03 of that network, then the name of the file would read as follows: N01_20200103_int_03_...

If, however, the interview was with a partner from an activist organization within that network, then the name would read N01_20200103_int_00_...

Device

A single event, involving the same network, date, type of data, and school number, may have been documented using multiple devices. For example, an interview may have been filmed, but also audiotaped. Device specifies the device used for recording in the local network. Each local network needs to name its own data collection devices, and keep track of and update the device name relation in their log relating codes defined within each local network (described in section 5.3). Typically, devices will be video or photo cameras and audio recorders, in which case the preferred naming would be VCAM# for video, FCAM# for photo, and AUD# for audio recorders. Of course, other more generic devices, such as notebooks, may have been used, or the source of data (a digital questionnaire) is already presumed in the type of data. However, to be consistent and make data as aligned with FAIR principles as possible, we expect each local network to keep track of a naming structure that also includes a code for device. If the Global Assessment Instrument is provided electronically, the code GAI (for Global Assessment Instrument) will be used.

Accordingly, if the data to be stored is an interview, recorded the third of January of 2020, in the Norwegian local network 1, with one of the teachers of school 03 of that network, but I have two files, recorded with two devices, VCAM1, and AUD1, then I will store two files as follows:

N01_20200103_int_03_VCAM1_...

N01_20200103_int_03_AUD1_...

File number

There are multiple situations in which the same event, recorded with the same device, will result in multiple files. Many video and audio recording devices generate multiple files rather than a large one, to keep sizes manageable. Also, recording may involve multiple start-stop situations in which new files are created automatically. When collecting documents (e.g., images, texts), we expect to be able to identify single files as they belong to single participants, for example, in which case we will prefer storing text and images separately rather than uploading a compiled file.

In all these cases, a file number is required. File numbers can be either copied from those automatically generated by the recording devices, or manually, selecting a digit number. In either case, when files have a chronological relation, it is required that file number will follow the chronological order. Thus, if the data to be stored is an interview, recorded the third of January of 2020, in the Norwegian local network 1, with one of the teachers of school 03 of that network, and I have two files generated chronologically with the VCAM1 device, the file will be named as follows, where file number 01 was chronologically prior to file number 02:

N01_20200103_int_03_VCAM1_01

N01_20200103_int_03_VCAM1_02

Format

A relation of suitable formats is detailed in section 4.4, particularly in table 1. Format does not need to be included in the name other than as the default format identifier inserted after a dot at the end of the file name. Most operative systems automatically identify a file format extension and include the right extension name at the end of the file. Thus, if the file in the example above was formatted as and Apple QuickTime Movie file (file extension .mov), the name of the resulting file would have been:

N01_20200103_int_03_VCAM1_01.mov

TSD archiving capabilities allow for cataloguing files with respect to type of file format, which shall allow users to find and access formats without need to insert a new element in the file name other than the extension.

5.4 Data log file(s)

One or more structured log files overseeing the entire database will be stored in the data log folder. The purpose of these files will be that users can seamlessly perform searchers (e.g., through keywords) to locate data relevant to their purposes. The number and specific structure of these files will be determined in collaboration by the local data managers and the overall data flow manager. Core information will include all aspects specified in the file name structure, expanded to provide more complete information, but also more detailed information concerning the precedents, conditions, and other contingent aspects relevant for researchers to navigate and use the data to achieve SEAS related goals. These files will be written as spreadsheets (e.g., Microsoft Office Excel files), and will be readable by all users. The information included in these files will be dependent on the information that local data managers provide as part of the data packages uploaded to the inbox folder. The overall data flow manager is in charge of updating the permanent data log files with the information provided by the local data log files.

5.5 Log relating codes defined within each local network

Because the number of schools/settings involved in each local network cannot be defined centrally, and the naming structure in SEAS requires using codes that identify these specific aspects, each local network is required to provide a log relating the codes defined within each local network as part of each data package uploaded. These logs, in turn, will be stored under the data log files folder.

5.6 Log file connecting data files and participants' consent.

A log file connecting data files with participants' consent forms and conditions will be generated and updated every time a new data package is delivered and stored. The log, a spreadsheet file, will contain information that shall make it possible for users to trace back files with respect to specific consent forms. In this way, (a) conditions of use for different files/participants are specified in one document, with the types of consent given being displayed for each participant; and (b) if conditions of use were to be changed, these can be updated, and eventual retrieval and deletion of files (e.g., if a participant would withdraw consent) will be possible.

5.7 Making data openly accessible and increase re-use

In principle, all raw data generated during the SEAS project shall contain diverse degrees of private information and will be treated confidentially, in line with approved ethics locally, and in line with GDPR and FAIR principles more globally. This means that, in principle, only consortium members, and users applying to access data will be granted permission to access the data stored in the data folder.

To access TSD and use SEAS data, consortium members need to sign a contract acknowledging that they will follow the instructions provided in this document.

In order to make the data accessible to members outside of the consortium, contractual conditions must be established regulating the purposes, access, and uses that the data can be given, to ensure compliance with the SEAS Consortium Agreement, the international ethical standards, GDPR, and the specifics of the consent given by participants. In this regard, once contractual conditions are established and agreed upon, users external to the consortium may be given access to TSD and follow the guidelines provided here. The University of Oslo has templates for such contractual agreements that shall be followed and tailored to fit SEAS requirements and purposes.

Processed data, on the other hand, shall be anonymized, and therefore accessible to users outside of the consortium. The Project Management Team (PMT) will suggest which sets of processed data can be published in a dedicated area for data sharing in the project's website, www.seas.uio.no, as well as will be responsible for identifying and tailoring specific contractual arrangements for accessing raw data. The Steering Committee (SC) of SEAS will decide on the suggestions promoted by the PMT. Processed data shall include anonymized transcriptions of interactions in relevant generic settings, aggregated data, and visualizations of quantitative as well as of qualitative analyses conducted as part of the project's local and global assessment activities. The processed and

anonymized data will also be shared by appropriate repository for long term storage after the project ends.

In addition to processed data being made available on the SEAS web site, and anonymised data being stored in an openly accessible repository after the project end, most deliverables in SEAS are built upon processed data and can themselves be considered and treated as aggregated data sets that will be made public either during the project period, or after the project period for deliverables to be kept confidential.

SEAS publications will adhere to green and gold open access publicing, preferably gold open access.

The data will be quality assured by the process described in section 4.3, and stored in TSD for minimum five years after project end.

5.8 Making data interoperable

The implementation of a FAIR access approach consisting of highly conventionalized shared folder structure, file name structure, and routines, as described here, opens for the possibility of making data interoperable across projects and research purposes connected to research interests within the area of open schooling and of scientific literacy and education for sustainability. In addition to the generic aspects structuring folders and name of files, categories relevant to the study of open schooling, scientific literacy, interest in science and other generic categories common to and relevant to these areas of research will be used to catalogue data through the data log files and folder. Coupled with the routines for making data openly accessible described above, and given the contractual conditions and GDPR are met, the SEAS data is, in principle, highly interoperable.

6 Allocation of resources

Costs for data storage in TSD is covered by the Horizon 2020 grant, and costs for open access publishing (green or gold) is similarly covered in partners' budgets. Partners will seek to maximize the extent of gold open access by utilizing additional sources of funding for such publicizing.

Coordinator is responsible for data management overall, and each local network has a Local Data Manager, responsible for implementing the DMP, and provide a local Data Management Plan in accordance with the SEAS overall DMP. WP3 leader oversees the processes of data importing, documenting, cleaning, and storing in the local network in the role of data flow manager.

Long term depository is planned to be serviced by the Norwegian Centre for Research Data (NSD) who is a certified infrastructure for archiving and sharing data[3]. This service is currently free of charge.

[3] https://nsd.no/arkivering/en/

7 Data security

The TSD - Service for Sensitive Data - is a platform which collects, stores, analyses and shares sensitive data in compliance with the Norwegian regulation on privacy. TSD is used by researchers working at UiO and in other public research institutions (UH-sector, university hospitals etc.). The aim of TSD is to offer storage and processing of data in a secure environment. To achieve a higher degree of security and to make working with control and risk management easier, TSD is disconnected from UiO's regular networks and offers just a limited number of ways one can interact with the system. All projects/user groups are hence issued with their own dedicated virtual network interconnecting any number of dedicated project servers.

SEAS will store personal sensitive data as students, including young students, will be identifiable on video recordings. In order to connect data constituting a learning trajectory or narrative, different data for individuals need to be connected with a key that traces participants as well as their consent (as described above in sections 5.2 and 5.3). For research purposes, each participant will be given a numerical code (to be used in place of a name). The technology should cater to the fact that each participant will be given a unique identification code, rather than a name, and all data will be securely stored and preserved, both electronically and on paper. Only authorized research personnel will have access to the password-protected electronic database. No unauthorized access will be possible. A separate data log file linking codes with names will be kept in a secure place. In order to safeguard the confidentiality of the participants' personal information, such data will be stored in a record that will be kept locked in the institution that acts as the leader of a local network in SEAS. Questionnaires and other paper materials collected in the project are stored in a locked cabinet in a secured room of the relevant institution with access restricted to authorized persons only. In order to guarantee open data for a consultation they will be stored for 5 years after the conclusion of the research project (2027) and will be subsequently destroyed. For future opportunities of research, data should be stored for 10 years after the conclusion of the project (2032), depending on the national regulations.

Data transfers to and from the services are handled by a special purpose file staging service and the project administrator (PC in SEAS) controls access rights for all project members. By default, all project members are able to transfer data in, but only the project administrator and the overall data flow manager (WP3 leader) can do a data transfer out. The data stored in TSD will be shared by their coded identity; the key relating number and true identity will not be included for data shared by selected researchers in SEAS.

8 Ethical aspects

8.1 Ethical requirements (Post-grant requirements) – WP 1

The post-grant requirement entails as part of WP 1 Ethical requirements:

- Details on incidental findings policy (regarding learning disadvantages etc.) must be submitted as a deliverable (D1.1 – due April 2020)
- Assent procedure must be explained and submitted as deliverable (D1.2 – due April 2020). This procedure is under development, cf. Appendix 1.

-The host institution must confirm in a deliverable that it has appointed a Data Protection Officer (DPO) and the contact details of the DPO are made available to all data subjects involved in the research. The DPO is senior adviser Roger Markgraf-Bye, personvernombud@uio.no and is announced at the consortium meeting in Tallinn, December 2019.

Appendix 1: Information and consent generic document

This document is a generic text for informing and obtaining consent for the acquisition of personal data in SEAS. The document is designed for adhering to GDPR in a Norwegian context, and to provide a template for a graded consent. However, this text must be adapted to the requirements, laws and regulations of the national context and to the settings of the local networks, including the age of the respondents.

Project information and consent <student/parents>

Invitation to take part in developing new teaching approaches to sustainable development

The current environmental challenges call for new developments in how science and other school subjects are taught, in order for students to develop important competencies for participation and action towards a sustainable society. What new forms of scientific knowledge, literacies, and skills do students and citizens need to address environmental challenges? And how can schools be part of sustainable development?

SEAS develops tools and methods that facilitate collaboration where teachers and students work together with out-of-school partners on solving sustainability challenges in their local communities through what we call "open schooling". Open schooling involves schools cooperating with other institutions and organizations to achieve community well-being.

SEAS will coordinate and investigate collaboration among six open schooling networks across Europe, including <*country*>. Besides <*country*> schools in <*name all other countries*>.also take part in SEAS. The overall aims of SEAS are:

(1) identifying core principles and best practices required for creating and sustaining open schooling collaborations, and

(2) promoting scientific literacies and skills necessary to engage in addressing real-life, complex sustainability challenges.

Together with school <name>, partners in SEAS will collaborate in a local network on the theme of <*some description of the theme of the local network*>. By taking part in the project, students will cooperate with partners <*short description of partners*>, and families will also be involved during the project period. In order to support and develop the work of local networks, and for creating new models of opens schooling and scientific literacy for sustainability, the project will collect and analyse data collected from selected activities in local networks. These data may be video recordings of selected classroom teaching, interviews, selected visits to locations in the local environment, and meetings with local partners. Log data from digital tools may also be collected, and students' products on paper will also be collected. The teacher may be asked to comment on students' products and video recordings of teaching in order to understand better how the lesson help students learn and how students' learning develops.

Participation in the research project requires consent from <*students and/or parents*>. We emphasize that participation in the research part of SEAS is fully voluntarily. The student may at any time withdraw the consent, and the students' data will be deleted. Students that do not participate in the research project, will take part in ordinary activities without any data being collected.

The data collected will be stored in a secure server, designed for storing sensitive data. The data storage is disconnected from the ordinary university network system and can be accessed only by limited number of people, following a set of pre-agreed procedures. Only researchers in SEAS registered for access to the secure server are allowed access. The data manager at <institution> is responsible that any deviance during data collection is duly registered, and that any irregularities during the collection are resolved before data are stored in the local server.
<The data is also stored in *IF LOCAL NETWORK ALSO STORES DATA LOCALLY IN ADDITION TO TSD, IT MUST BE DESCRIBED*>.

Dissemination of SEAS data

Dissemination means that data may be used in dissemination from the project, for research or general public dissemination. The general rule is that all students will be anonymous - not recognizable - in any dissemination. However, we will ask the student if he/she allows for use of data in presentations on web and social media, and for teaching purposes, where face or voice is recognizable. We emphasize that these additional choices are fully optional.

The student/child takes part with full anonymity if the additional options are not selected.

Audio-visual data on which the student can be identified will be deleted in 2032 <if permitted. 5 years after project ends in 2022 is minimum, we should opt for 10>. Any other data will be anonymised and delivered to the repository for long term storage after storage time period.

Please observe the consent form on page 3.

By making contact with <*Local Network Leader*> you are entitled to:

- Get access to personal information stored about the student
- Ensure personal information is deleted
- Get a copy of own data (data portability).

You may also file a complaint regarding any misuse of your personal information to the <National authority. In Norway: Norwegian Centre for Research Data (NSD) or The Norwegian Data Protection Authority (DPA)>.

The data management in SEAS is governed by strict regulations according to the General Data Protection Regulations (GDPR) and national regulations. For more information, please contact <*local network responsible*>.

Declaration of Consent

I have read the above information regarding the SEAS project, and understand that the voluntary participation in the project implies the collection of audio-visual data, pictures, log file data from digital tools, and student products.

1. Consent for participation in the SEAS research project. Please mark one of the options:
   My/our child takes <and I/we if parents take part> take part in the SEAS research project.
   - o Yes, I/we consent.
   - o No, I/we do not consent

2. Consent for collecting audio-visual materials

> By audio-visual materials are meant pictures, soud- or video recordings where voices may be recognizeable and faces may be recognizeable. It also includes student products such as drawings or exhibition items. Name of student will not be connected to the material.

Please provide information on the use of audio-visual materials in dissemination. Please mark the choices that apply for your child:

Audio-visual materials can be used for illustrative purposes in lectures, teaching and research articles, providing good examples of practices or learning experiences. Materials can be made available on digital platforms with restricted access in teacher education programmes.

- o Yes, I/we consent.
- o No, I/we do not consent

Audio-visual materials can be used in presentations and articles in open access media.

- o Yes, I/we consent.
- o No, I/we do not consent

Name of the student: _____

Parents signature: _____

Place and date: _____



SEAS

Science Education for Action and engagement towards Sustainability

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824522.



UiO : University of Oslo