



DELIVERABLE 1.5
DATA STORAGE REPOSITORY
REPORT

31 March 2022



Summary Sheet

Deliverable Number	D1.5
Deliverable Name	Data storage Repository Plan
Full Project Title	SESA – Smart Energy Solutions for Africa
Responsible Author(s)	Esther van Bergen; Daan van Rooij; Jorden van der Hoogt (CENEX NL)
Contributing Partner(s)	TUB, ICLEI AS, DTU
Peer Review	ICLEI ES, BTH, TEC, UNEP, F6S
Contractual Delivery Date	31-03-2022
Actual Delivery Date	31-03-2022
Status	Final Version
Dissemination level	Confidential, only for members of the consortium (including the Commission Services)
Version	V1.0
No. of Pages	24
WP/Task related to the deliverable	WP 1 / T 1.3
WP/Task responsible	CENEX NL
Document ID	SESA_D1.5_Data Repository Plan
Abstract	Data is both generated and used to perform tasks for the development of content for the SESA Toolbox (WP1/T1.1/D1.1) and for the SESA Evaluation Framework (WP1/T1.2/D1.2). This encompasses activities (to be) performed in relation to the energy solutions designed and implemented within the project at demonstration, validation and replication locations. This document summarizes the approach description and structure for the data storage repository for these tasks.

Legal Disclaimer

SESA (Grant Agreement No 101037141) is an Innovation Action project funded by the EU Framework Programme Horizon 2020. This document contains information about SESA core activities, findings, and outcomes. The content of this publication is the sole responsibility of the SESA consortium and cannot be considered to reflect the views of the European Commission.

Table of contents

List of Figures	3
List of Tables	3
List of Abbreviations	3
Executive Summary	4
1. Introduction to this report	5
1.1. About the SESA project - Smart Energy Solution for Africa	5
2. SESA data for Toolbox and Evaluation	6
2.1. Data repository definition	6
2.1.1. Data repository in H2020	6
2.1.2. Data repository in SESA WP1	6
2.2. Expected data	7
2.2.1. Data types and forms	7
2.2.2. Data sensitivity	8
2.2.3. General data example	8
2.3. Data flows	9
2.3.1. Input data	9
2.3.2. Output data	10
2.4. Requirements	10
2.4.1. Governance structure	12
2.4.2. Access and security	13
2.4.3. Data type storage	14
2.4.4. Metadata	14
2.4.5. Archiving and long-term preservation	14
2.4.6. Ease of use	14
2.4.7. Availability	14
2.4.8. Established project environment	14
2.4.9. Costs	14
2.5. Conclusion	14
3. Repositories for Work Package 1	16
3.1. Evaluation and Assessments Repository	16
3.2. Sign-off Portal Repository	17
3.3. Back-end Toolbox Repository	18
4. Platform Selection	19
4.1. Repository long list	19
4.2. Evaluation and Assessments Repository	20

4.3. Sign-off Portal Repository	20
4.4. Back-end Toolbox Repository	21
4.5. Implementation – Next steps.....	21

List of Figures

Figure 1: Overview of data flows within WP1 repository environment.....	9
Figure 2: Overview of the three repositories needed in WP1	15

List of Tables

Table 1: Criteria for SESA repository environment	11
Table 2: Evaluation and Assessments Repository requirements	16
Table 3: Sign-off Portal Repository requirements	17
Table 4: Back-end Toolbox Repository requirements	18
Table 5: Repository long list	19
Table 6: SharePoint limits by plan	21

List of Abbreviations

<i>Abbreviation</i>	<i>Definition description</i>
<i>AGA</i>	Annotated Model Grant Agreement; a detailed document produced by the European Commission explain to applicants and beneficiaries the General Model Grant Agreement.
<i>DOI</i>	Digital Object Identifier; technical and social infrastructure for the registration and use of persistent interoperable identifiers for use on digital networks
<i>DCP</i>	Data Collection Plan; Deliverable D1.4 of the SESA project
<i>DMP</i>	Data Management Plan; Deliverable D7.5 of the SESA project
<i>DRP</i>	Data storage Repository Plan; Deliverable D1.5 of the SESA project (this report)
<i>GA</i>	Reference to the SESA specific Grant Agreement
<i>GDPR</i>	General Data Protection Regulation; Europe’s data privacy and security legislation
<i>H2020</i>	European Union’s Horizon 2020 research and innovation programme
<i>KPI</i>	Key Performance Indicator; means to measure success or results in relation to the perceived objectives
<i>SESA</i>	Sustainable Energy Solutions for Africa; European Union’s Horizon 2020 research and innovation programme under Grant Agreement No. 101037141.
<i>WP</i>	Work Package; A major sub-division of the (Horizon 2020) project which includes several key tasks and may have specific Milestones and / or Deliverables as output-results.

Executive Summary

SESA – Smart Energy Solutions for Africa is a collaborative project between the European Union and nine African countries that aims at providing innovative energy solutions using decentralized renewables. This Data storage Repository Plan (DRP) report describes the approach for its repository environment based on how the data flows in relation to the development of the SESA Toolbox and the Evaluation Framework. This relates to WP1 itself as well as the data generated and used for activities in other WPs, both for the creation of content for the Toolbox and the Evaluation Framework's various analysis and assessments about the energy solutions that are demonstrated, validated and replicated.

This repository environment is not meant to function as an open-science database and access is therefore intended to remain internal to SESA partners. Data intended to be made publicly available may flow through this environment before it will be shared via SESA's publicly accessible dissemination channels and activities.

At the time of publication of this report the exact data stored in this repository environment is not yet defined. However, we can expect this data to be both quantitative and qualitative in nature and can contain raw, processed and (developed from this) output data. Additionally, the data could potentially be structured, semi-structured or even unstructured. Data can also vary in their levels of sensitivity. Data stored in this repository environment will be labelled with a sensitivity-indicator to ensure access permissions can be arranged accordingly. This means the more sensitive the data stored in the repository environment, the better access and security requirements will need to be for the selected platform. In addition to access and security, seven other 'requirements categories' were used for the selection process of the relevant repository platform.

There are three main but different data flows identified for the Data storage Repository environment. The conclusion is these are best served by three separate repositories. From a long-list of available platforms, the best candidates were identified based on the 'requirements categories' and a platform suggestion ensued. Final decisions are to be confirmed by the Project's Steering Group. Implementation will follow once information is available of certain outstanding questions as well as more detailed insights for the process of data collection.