Lithium niobate empowered silicon nitride platform for fragmentation-free operation in the visible and the NIR

Funded by the European Union

Deliverable D7.2 – Data management plan

Lead Beneficiary	ICCS
Contact Person	Lefteris Gounaridis
Address	9, Iroon Polytechneiou, 15780 Athens, GREECE,
Phone	+30 210 772 2057
e-mail	lgou@mail.ntua.gr
Date due of deliverable	28.02.2023 [M06]
Actual submission date	09.06.2023 [M10]
Authors	Lefteris Gounaridis, Costis Christogiannis
Participants	ICCS
Work-package	WP7
Dissemination level	PU
Type	DMP
Version	1.0
Total number of pages	27

Document History

Version	Date DD.MM.YY	From	To	Description
1.0	09.06.23	ICCS	ICCS	Version 1.0 of the Deliverable

Copyright Statement

Funding Acknowledgement

Table of Contents

List of abbreviations

APCs	Author Processing Charges
CA	Consortium Agreement
DMP	Data Management Plan
DOI	Digital Object Identifier
DSP	Digital Signal Processing
EC	European Commission
ECL	External Cavity Laser
EU	European Union
EOSC	European Open Science Cloud
FAIR	Findable, Accessible, Interoperable, Re-usable
GA	Grant Agreement
HTTP	Hypertext Transfer Protocol Secure
IA	Innovation Action
ICT	Information and Communication Technology
TIMT	Technical and Innovation Team
IPR	Intellectual Property Rights
LNOI	Lithium Niobate On Insulator
MZM	Mach-Zehnder modulator
OA	Open Access
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting
PC	Project Coordinator
PCB	Printed Circuit Board
PIC	Photonic Integrated Circuit
PtP	point-to-point
RAID	Redundant Array of Independent Disks

Executive Summary

This document provides the initial version of the Data Management Plan (DMP) to be implemented by the LOLIPOP project team. The LOLIPOP DMP is a deliverable of WP7 and is a document that prescribes how LOLIPOP contributes to the Horizon Europe objectives for Open Research Data Pilot, introduced in aiming to improve and maximize access to and re-use of research data generated by projects. In line with this decision, in this initial phase LOLIPOP has identified several data sets which will be generated within the project and shared with the research community. The main objectives of the project are to improve the exchange and dissemination of research results and possibly to enable and promote a wider validation of the project results and to encourage a fair comparison and evaluation of different solutions in the technical areas of LOLIPOP.

The Data Management Plan deals with all the concerns about the treatment of the data that will be collected, processed, or generated in the whole project lifecycle and has been structured in compliance with the guidelines and the template conveyed by the European Commission.

Hereafter are reported the main aspects that will be considered in the Data Management Plan, for each of the data set identified in the project:

• Types of data generated, collected, or processed

• Standards used to manage data

• Data exploitation methodology

• Accessibility to data produced by the Project

• Data Dissemination level

• Data Preservation and re-use

The DMP is a 'live' document in which the respective information will be made available on a finer level of granularity through updates as the implementation of the project progresses and when significant changes occur. This document is the first version of the DMP, providing an initial description of the expected data sets and their management plan, while more details and potentially further data sets might be added during the project implementation. The document will be maintained active and regularly updated with additional data sets or details about the existing ones during the total duration of the project.

Keywords: Data Management Plan (DMP), open science, FAIR (Findable, Accessible, Interoperable, Re-usable data), Dissemination level, Data sets, research data management.

Definition and Terminology

Dataset: Digital information created during research, but which is not a published research output. Research data excludes purely administrative records. The highest priority research data is that which underpins a research output. Research data does not include publications, articles, lectures, or presentations.

Background: any data, know-how or information — whatever its form or nature (tangible or intangible), including any rights such as intellectual property rights — that is:

(a) held by the beneficiaries before they acceded to the Agreement and

(b) needed to implement the action or exploit the results.

Results: any tangible or intangible effect of the action, such as data, know-how or information, whatever its form or nature, whether it can be protected, as well as any rights attached to it, including intellectual property rights.

Data Management Plan (DMP): A formal working document which outlines how datasets will be handled both during the active research phase and after the project is completed. DMPs in some forms are now a requirement of a research project and therefore must be addressed at the earliest phase of the research lifecycle.

Digital Object Identifier (DOI): A persistent identifier or handle used to identify objects uniquely and to direct communications to the correct server, developed by the International DOI Foundation. An implementation of the Handle System, DOIs are in wide use mainly to identify academic, professional, and government information, such as journal articles, research reports and data sets, and official publications.

Metadata: Information about datasets stored in a repository/database template. For example, an image may include metadata that describes how large the picture is, the color depth, the image resolution, when the image was created, and other data. A text document's metadata may contain information about how long the document is, who the author is, when the document was written, and a short summary of the document.

Repository: A digital repository is a mechanism for managing and storing digital content. Repositories can be subject or institutional in their focus.

Zenodo: A general-purpose open-access repository (https://about.zenodo.org) developed under the European OpenAIRE program and operated by CERN. It allows researchers to deposit data sets, research software, reports, and any other research related digital artifacts. For each submission, a persistent digital object identifier (DOI) is minted, which makes the stored items easily citable.

1   Introduction

The Horizon Europe framework recommends that a data management plan ('DMP') is established and regularly updated. The scope of the DMP is to describe how to select, structure, store and make FAIR (Findable, Accessible, Interoperable and Re-usable), the "background" information used, and the project "results" generated during the project.

The present deliverable is the first release of the LOLIPOP DMP, and it is structured along the guidelines of Data Management Plan (HE) template. In completing the sections of the template, the requirements on "Intellectual Property Rights- Background and Results-Access Rights and Rights of Use" as stipulated in Article 16, and on "Communication, Dissemination and Visibility" as stipulated in Article 17 in the Annotated Grant Agreement, are considered.

Horizon Europe framework has adopted the Open Science approach in its Data Management strategy - that is based on open cooperative work and systematic sharing of knowledge and tools as early and widely as possible in the process. The Open Science approach aims at increasing the quality and efficiency of research and accelerate the advancement of knowledge and innovation by sharing results, making them more reusable and improving their reproducibility. Open Science and Open Access requirements and how LOLIPOP fulfills them are discussed in this document.

The DMP contains information related to the types of data the project will collect, use, and generate, the data standards to be employed and how project partners may disseminate and exploit project results. The deliverable also reports on the data sets that will be made available in open repositories together with the necessary reference details for access to the open data.

In particular, the DMP identifies:

♦   What types of data will be generated or collected.

♦   What data will be exploited/and which will be the data dissemination level.

♦   What standards will be used to manage data/ metadata.

♦   How will data be preserved, including after project completion.

This document provides an initial description of the expected data sets and their management plan, while more details and potentially further data sets might be added during the project implementation. The document will be maintained active and regularly updated with additional data sets or updates on the existing ones. As a minimum, updated versions of the DMP document will be released with D7.6 - Interim exploitation and dissemination report and updated DMP, due in M30 and D7.8 - Final exploitation and dissemination report and updated DMP, due in M42.

2   LOLIPOP – Project Overview

LOLIPOP is a 42-month Research Programme part of the HORIZON Innovation Actions funded by the European Union through the **Horizon Europe** framework. It is launched on 1st September 2022.

Today, silicon nitride in photonics is gaining tremendous speed. The silicon nitride is probably the platform with the highest potential for the development of complex photonic integrated circuits (PICs), which can sense and think with ultra-high sensitivity, precision, and computational resources. This is due to the combination of low propagation loss (below 0.1 dB/cm), high integration density, and ultra-wideband operation that it can offer. This upcoming generation of PICs could increase the use of photonics in already-existing application fields like metrology, sensing, biosensing, and microwave photonics as well as make it possible to adopt photonic solutions in newly emerging fields with high interest like neuromorphic computing, quantum communications, and quantum computing.

The silicon nitride platform is passive despite these great prospects and its special potential for passive waveguiding performance. On this platform, hybrid solutions for light emission, modulation, nonlinear processing, and detection are currently poor or absent if they are to simultaneously provide high photonic performance, wideband operation, and integration robustness. Thus, the silicon nitride platform has potential, but it is only partially being utilized currently. LOLIPOP aims to close this gap so that the silicon nitride platform can advance and develop fully. To achieve this, LOLIPOP invests in the combination of silicon nitride with the lithium niobate on insulator (LNOI) technology while concurrently utilizing a disruptive set of techniques to integrate semiconducting materials onto the silicon nitride substrate.

Despite huge progress in photonics, extended spectral bands at wavelengths below 1100 nm remain heavily underserved in terms of integration solutions. At the same time, silicon nitride is booming, and lithium niobate is making an impressive comeback in the form of lithium niobate on insulator (LNOI), with both materials being transparent both in the visible spectrum and near infrared (NIR) part of the spectrum. With all these viewed as a unique opportunity, LOLIPOP steps in to develop a disruptive platform that will offer the highest integration, modulation, and second order nonlinear performance in the entire spectrum from 400 nm up to 1600 nm, based on the combination of LNOI and silicon-nitride (TriPleX®) technology. To this end, LOLIPOP will develop die-bonding method for the integration of the gain elements on TriPleX, and micro-transfer-printing method for low-loss (<0.5 dB) integration of LNOI films on TriPleX® without compromising in functionality of the two platforms.

LOLIPOP will also develop a process for the growth of Ge photodiodes (PDs) inside pockets and a process for the flip-chip bonding of active elements inside recesses within TriPleX®. Given the possibility of the Ge-PDs to operate in the entire 400 -1600 nm spectral area, and the flexibility of the bonding process to adapt to different actives and wavelengths, the picture of this ultra-wideband technology is complete.

3   LOLIPOP Data Management Planning

The present section describes the plans of LOLIPOP consortium for the management of the different sets of research data that will be produced, collected and used for internal processing and validation in the project.

3.1   Open Science by LOLIPOP

LOLIPOP will apply the Open science policies as prioritized by the Horizon Europe framework. Open science practices will be applied to ensure that research data, of all sorts, generated in the course of the project is Findable, Accessible, Interoperable, and Reusable (FAIR). This means that interested parties – academia, industry, end users, civil society will be able to search for, find and access research as well as to understand and use the research data. This access will help in maintaining and checking the quality of research conducted in the framework of the project, but also will offer an opportunity to interact with interested communities.

Open Science Practices

Open science practices include early and open sharing of research (for example through preregistration, registered reports, pre-prints, or crowdsourcing); research output management; measures to ensure reproducibility of research outputs; providing open access to research outputs (such as publications, data, software, models, algorithms, and workflows); participation in open peer-review.

In the framework of Horizon Europe, a couple of open science practices are identified as mandatory, such as:

(i)   open access to scientific publications,

(ii)   responsible management of research data in line with the FAIR principles of 'Findability', 'Accessibility', 'Interoperability' and 'Reusability', under the principle 'as open as possible, as closed as necessary',

(iii)   information about the research outputs/tools/instruments needed to validate the conclusions of scientific publications or to validate/re-use research data,

(iv)   digital or physical access to the results needed to validate the conclusions of scientific publications, unless exceptions apply,

(v)   in cases of public emergency, if requested by the granting authority, immediate open access to all research outputs under open licenses or, if exceptions apply, access under fair and reasonable conditions to legal entities that need the research outputs to address the public emergency.

However, due to IP or commercial considerations, may not all data/tools/methods be at open access. Potential restrictions to open access will be evaluated on a case-by-case basis. In such cases – where immediate open access is not recommended, the consortium is committed to depositing research data and outputs retrospectively in repositories and provide open access at the earliest possible.

In the remaining section, the way of how those practices is applied by LOLIPOP is discussed.

3.1.1 Open access to publications

The first aspect to be considered in the DMP is related to the open access (OA) to the publications generated within the context of the project and made available online to any user at no charge. In line with the Horizon Europe guidelines on open access to scientific publications, the publications that will arise from the LOLIPOP project will be made public through the following two channels:

a) **Open access publishing** (also referred to as **'Gold' open access**) means that an article is immediately placed in open access mode (on the publisher/journal website). In this option, beneficiaries publish their results in open access journals, or in hybrid journals that sell subscriptions and offer the possibility of making individual articles openly accessible via the payment of Author Processing Charges (APCs). LOLIPOP will preferably target 'Gold' open access journals which apply article processing charges for the article to be published and then it is freely available online via the Open Access journal site.

b) **Self-archiving** (also referred to as **'green' open access**). The project will use widely self-archiving (or green open access) services for research community like ResearchGate or Academia, that will allow balance between traditional publications and open-access. In this option, the beneficiaries deposit the final peer-reviewed manuscript in an online repository of their choice (e.g., Zenodo). If this route is chosen, the beneficiaries will ensure open access to the publication within a maximum of six months, in line with the open access obligations established by the EC.

The consortium partners will use the Open Research Europe publishing platform to enable direct publication of the research outputs in support of research integrity, reproducibility, transparency, and open science practices. We should note here that there may be cases where the project participants may need to submit articles to journals (or proceedings) that only offer a lower level of open access, requiring either parallel publication or an embargo period.

Additionally, whenever possible the LOLIPOP consortium will retain the ownership of the copyright for their work using a 'License to Publish', which is a publishing agreement between author and publisher. With this agreement, authors can retain copyright and the right to deposit the article in an Open Access repository, while providing the publisher with the necessary rights to publish the article.

All publications will acknowledge the project funding. This acknowledgment will be included also in the metadata of the generated information since it allows to maximize the discoverability and visibility of publications and to ensure the acknowledgment of EU funding.

Metadata of deposited publications will be open under a Creative Common Public Domain Dedication (CC 0) or equivalent, in line with the FAIR principles (in particular machine actionable) and provide information at least about the following: publication (author(s), title, date of publication, publication venue); Horizon Europe ; grant project name, acronym and number; licensing terms; persistent identifiers for the publication, the authors involved in the action and, if possible, for their organizations and the grant. Where applicable, the metadata will include persistent identifiers for

any research output, or any other tools and instruments needed to validate the conclusions of the publication (GA, Article 17)

To ensure open access — via the repository — to the bibliographic metadata that identify the deposited publication, the metadata must be in a standard format and will include all the following:

- the terms "European Union (EU)" and "Horizon Europe"
- the Name of the action, the Acronym and Grant Number
- the publication date and length of embargo period (if applicable), and
- a persistent identifier (e.g. DOI)

3.1.2   Research Data Management

Apart from the open access to publication explained in the previous section, LOLIPOP will adopt the European Open Science Cloud (EOSC) as the trusted environment for hosting and processing research data. The EOSC portal capabilities will be explored and embraced as a medium to underpin the open access research policy. In specific, two types of data sets are defined:

Background: any data, know-how or information (tangible or intangible), including any rights such as intellectual property rights, that already exist or have been already generated and reported by other research initiatives or held by the consortium members from their own research and development activities prior or in parallel with the project and which might be used as inputs in LOLIPOP in implementing the project or exploiting its results. LOLIPOP will act only as recipient of this type of data and, if just re-used "as-is", the project will not define any specific strategy for their management which is already handled by other organizations.

Results: any data, know-how (tangible or intangible) - whether it can be protected, as well as any rights attached to it, including IPRs - that is produced because of project activities.

Before a dataset is deposited in an open access repository, the partner(s) responsible for uploading and preserving the data will seek consent for data preservation and sharing from all partners involved (according to the provisions of the CA and with the guidance of the Technical and Innovation Management Committee [TIMT]). The respective partners, in collaboration with the Project Coordinator (PC) and the Technical and Innovation Manager (TIM) will be responsible for postponing or restricting data sharing to allow enough time for publishing the results in peer-reviewed journals or for seeking patents. To facilitate handling of datasets especially where multiple partners are involved, to expedite data dissemination and ensure that no ethical issues are associated to the respective datasets, the process will be overseen by the TIMT.

3.2   FAIR (Findable, Accessible, Interoperable, and Reusable) data

Data and research outcomes generated by LOLIPOP activities will be managed and curated in line with the FAIR principles. In summary, the application of FAIR principles for LOLIPOP generated data, i.e., publications, simulation data and experimental data as identified in DoA, is as follows:

For Publications

F: Reputable publications, A: Green open access or gold open access; deposited in trusted repositories e.g., Zenodo,

I: Readable online, downloadable, and printable, Machine-readable copies, standard text file formats. Use of bibliographic metadata including terms, name of project, acronym and grant number, publication data and persistent identifier, **R**: Efforts will be made to provide right to copy, distribute, search, link, crawl, and mine on an as-needed basis. (**Curation**: By the author, Article Processing Charges (APCs) for gold open access have been included in partners' other costs where applicable).

For Simulation Data from components/sub-system modelling activities

F: Types of persistent, unique identifiers (e.g., digital object identifiers); deposited in trusted repositories e.g., Zenodo,

A: Open access over standard protocols (e.g., HTTP and OAI-PMH); IPR considerations and timeline for open access; provisions for access to restricted data for commercial exploitation and/or verification purposes,

I: Standards, formats and vocabularies for data and metadata; accompanied with a "README" file describing any of the relevant information to a specific dataset,

R: Licenses for data sharing and re-use (e.g., Creative Commons, Open Data Commons); availability of tools/software/models for data generation and validation/interpretation/reuse. Storage/preservation costs; person/team responsible for data management and quality assurance.

For Experimental data from components/sub-system characterization and demonstrators testing

F: Types of persistent and unique identifiers (e.g., digital object identifiers); deposited in trusted repositories e.g., Zenodo

A: Open access over standard protocols (e.g., HTTP and OAI-PMH); IPR considerations and timeline for open access; provisions for access to restricted data for commercial exploitation and/or verification purposes,

I: Standards, formats and vocabularies for data and metadata; accompanied with a "README" file describing any of the relevant information to a specific dataset,

R: Licenses for data sharing and re-use (e.g., Creative Commons, Open Data Commons); availability of tools/software/models for data generation and validation/interpretation /re-use. Storage/preservation costs; person/team responsible for data management and quality assurance.

3.2.1   Making data findable, including provisions for metadata

For the open datasets, a Digital Object Identifier (DOI) will be assigned to datasets for effective and persistent citation when it is uploaded to the repository [Zenodo]. This DOI can be used in any relevant publications to direct readers to the underlying dataset.

Each dataset generated during the project will be recorded in an Excel spreadsheet with a standard format and allocated a dataset identifier according to Table 1. The owner of each dataset will be responsible for the creation of the spreadsheet associated with the generated dataset, while ICCS (Project Coordinator) will be responsible for gathering and preserving the information of all the dataset in its local repository.

LOLIPOP naming convention for project datasets will comprise of the following:

- A unique chronological number of the datasets in the project will be added.
- The title of the dataset.
- A prefix "LOLIPOP" indicates an LOLIPOP dataset.
- A unique identification number linking with the dataset work package e.g., "WP2".
- For each new version of a dataset, it will be allocated with a version number which will be for example start at v1.0.

Search keywords will be provided when the dataset is uploaded to Zenodo, which will optimize possibilities for re-use. Zenodo follows the minimum Data Cite metadata standards. The information and metadata stored in the record-spreadsheet for each generated dataset is summarized in Table 1 below.

Table 1. The structure (fields) of the record-spreadsheet for each generated dataset in LOLIPOP

Data set Reference Name	*The reference name will be based on the naming convention outlined above in Paragraph 3.2.1*
Description	*Description of the Dataset*
Standards and Metadata	In the absence of a well-defined metadata standard for this type of data, a simple README file will be used. This will be generated in raw text format and will describe basic details that will help people to find the data, including who created or contributed to the data, its title, date of creation and under what conditions it can be accessed. Documentation will also include details on the methodology used as well as file and folder naming conventions. The following fields will be used:
	1. **Dataset Title** *The title of the dataset which should be easily searchable & findable*
	2. **Name(s) of dataset creator(s)** *Lead partners responsible for the creation of the dataset*
	3. **Description of Data** *Brief description of the open data to be included in the metadata*
	4. **Data Source** *How/why was the dataset generated.*
	5. **Creation Date** *Date of generation of the dataset*
	6. **Format** *Possible formats of the datasets e.g., DOC, XLSX, PDF, JPEG, TIFF, etc.*
	7. **Expected size** Approximate size of the dataset
	8. **Location of Data** Institutional repository where the data are stored.
	9. **Digital Object Identified (DOI)** *The DOI can be entered once the dataset has been deposited in the repository*
	10. **Access status** Type of Dataset "Open" or "Restricted"

Data set Reference Name	*The reference name will be based on the naming convention outlined above in Paragraph 3.2.1*
	11. Embargo Embargo period of the dataset (if applicable) **12. Funding Statement** This project has received funding from the European Union's Horizon Europe Framework Programme (HORIZON) under G.A No 101070441. The results of this dataset reflect only the creator's view and the Commission is not responsible for any use that may be made of the information it contains. **13. Work Package** *LOLIPOP Work Package associated with this dataset* **14. Related publications** Bibliographical details of publications based on the dataset will be listed, with links to abstracts and, where possible, full text. **15. Dataset Citation** A 'ready-to-use' citation reference for the dataset will be provided – incorporating the core descriptive elements. **16. Keywords** Dataset related Keywords **17. Version number** Dataset version number in order to keep track of changes to the dataset.
Data sharing	♦ Repository Expected repository to be submitted. ♦ Date of Repository Submission *The date of submission to the repository will be added once it has been submitted*
Archiving and preservation	For redundancy, besides uploading the data on Zenodo it will be also maintained on a university-owned storage server at ICCS (PCRL group). The server offers real-time data mirroring through RAID (redundant array of independent disks) and weekly backups to external disk drives. The dataset will be preserved for at least 5 years after the project end and the associated costs will be covered by ICCS through own funds.

3.2.2 Making data openly accessible

Research data that is created during the project is owned by the beneficiary that generates them (G.A. Article 16). Each beneficiary must disseminate its results as soon as possible unless there is legitimate interest to protect the results. A beneficiary that intends to disseminate its results must give advance notice to the other beneficiaries — unless agreed otherwise — at least 45 days, together with enough information on the results it will disseminate (G.A. Article 17).

Generated research data will be deposited as follows:

⇨ **_Restricted data_** will be deposited in the repository of the partner that owns it.

⇨ ***Data accessible by the consortium members will only*** be deposited in the SharePoint Documents Repository of LOLIPOP (Private area for consortium members only). More specifically a folder(-s) will be set up in the project's private area, for easy upload of project datasets visible to the consortium members.

⇨ ***Open data*** will be deposited in Zenodo repository. Zenodo.org is open, free, searchable, and structured with flexible licensing allowing for storing all types of data: datasets, images, presentations, publications and software. In addition, Zenodo repository allows researchers to deposit both publications and data, while providing tools to link them.

LOLIPOP project has chosen to use Zenodo.org as the repository for storing the open project data for the following reasons:

- ✓ It enables *Shared Research* through the reposition of all research outputs from across all fields of research and science.
- ✓ It is *Citable and Discoverable:* uploads get a Digital Object Identifier (DOI) to make them easily and uniquely citable.
- ✓ It *fosters the establishment and curation of research Communities* through the creation of digital repositories, fully controlled by the owner.
- ✓ Allows the *identification of grants*, integrated in reporting lines for research funded by the European Commission via OpenAIRE.
- ✓ Allows for *Flexible licensing* and,
- ✓ *Ensures security of the stored research results* for the future in the same cloud infrastructure as CERN's own LHC research data.

Data objects will be deposited in Zenodo under:

- Open access to data files and metadata and data files provided over standard protocols such as HTTP and OAI-PMH.
- Use and reuse of data permitted.
- Privacy of its users protected.

For the data being deposited in an external repository (i.e., Zenodo), a dataset registry record will also be created in LOLIPOP Documents Repository. The registry record will be updated by the partner that is responsible for the specific generated data and will include relevant metadata explaining what data exists and a DOI link to where the data is available in the external repository.

During embargo periods, information about the restricted data will be published in the data repository, and details of when the data will become available will be included in the metadata. Where a restriction on open access to data is necessary, attempts will be made to make data available under controlled conditions to other individual researchers. In accordance with G.A. Article 25, data must be made available to partners upon request, including in the context of checks, reviews, audits, or investigations. Data will be made accessible and available for re-use and secondary analysis.

In parallel to the available Open Research Datasets, several restricted datasets will be generated within the project. These datasets that contain critical details on LOLIPOP developments will be protected in a restricted area. Each LOLIPOP partner is responsible for identifying these datasets, to

communicate with the consortium the updates on these datasets. Of course, these datasets can be moved from restricted areas to the Open Data zone of LOLIPOP if their owners select to partially or fully release them. This change will be resembled in the Data Management Plan of the project.

3.2.3   Making data interoperable

The LOLIPOP project aims to collect and document the data in a standardized way to ensure that the datasets can be understood, interpreted and shared in isolation alongside accompanying metadata and documentation.

Generated data will be preserved either in the LOLIPOP SharePoint Documents Repository or in the institutional intranet platforms until the end of the Project (see 4).

a) **General Information**
  - Title of the dataset
  - Dataset Identifier
  - Responsible Partner
  - Author Information
  - Date of data collection
  - Geographic location of data collection
  - Title of project and Funding sources that supported the collection of the data.

b) **Sharing/Access Information**
  - Licenses/access restrictions placed on the data.
  - Link to Data Repository
  - Links to other publicly accessible locations of the data
  - Links to publications that cite or use the data.
  - If the data derived from another source?

c) **Dataset/File Overview**
  - This dataset contains X sub-dataset as listed below:…………………………
  - What is the status of the documented data? – "complete", "in progress", or "planned."
  - Are there plans to update the data?

d) **Methodological Information**
  - Description of methods used for experimental design and data collection: *<Include links or references to publications or other documentation containing experimental design or protocols used in data collection>*
  - Methods for processing the data: *<Describe how the submitted data were generated from the raw or collected data>*
  - Instruments and software used in data collection and processing-specific information needed to interpret the data
  - Standards and calibration information, if appropriate
  - Environmental/experimental conditions
  - Describe any quality-assurance procedures performed on the data
  - Dataset benefits

An indicative example of a metadata file can be found in Appendix I of the present document.

3.2.4 Increase data re-use (through clarifying licenses)

The open datasets will be made available for re-use through uploads to the Zenodo community page for the project. In principle, the data will be stored in Zenodo after the conclusion of the Project without additional cost. All the research data will be of the highest quality, have long-term validity and will be well documented for other researchers to be able to get access and understand them after 5 years.

If datasets are updated, the partner that possesses the data has the responsibility to manage the different versions and to make sure that the latest version is available in the case of publicly available data. Quality control of the data is the responsibility of the relevant responsible partner generating the data.

4 Data Summary

LOLIPOP will produce several datasets during the lifetime of the project. The data will be both quantitative and qualitative in nature and will be analyzed from a range of methodological perspectives for project development and scientific purposes. These will be available in a variety of easily accessible formats, including Postscript (PDF, XPS), Excel (XLSX, XLS, CSV), Word (DOCX, DOC, RTF), Power Point (PPTX, PPT), image (JPEG, PNG, GIF, TIFF), compressed formats (TAR.GZ, MTZ).

Table 2 below summarizes the data that is foreseen to be generated during the implementation of LOLIPOP project. They are listed under each of the Work Package taken from the GA Annex 1 – Description of Action (DoA). The datasets will have the same structure, in accordance with the guide of Horizon Europe for Data Management.

Table 2. The potential data that will be generated in LOLIPOP project.

Data Description	Related Deliverable(s)	Type	Format	Estimated Volume	Access	IPR Owner
WP2. Design of LOLIPOP technology and demonstrators and development of control software						
Simulation studies of LDV modules The data set will include simulation results regarding the system performance of the LDV modules (Module-1 and -2).	D2.3	Results	TXT, TSV, CSV, JPEG	<500 MB	Open	1. ICCS
Simulation studies of Squeezed State Source module The data set will include simulation results regarding the impact of the insertion losses and crosstalk on the system performance of Module-3.	D2.3	Results	TXT, TSV, CSV, JPEG	<500 MB	Open	10. QuiX
Simulation studies of LiDAR module The data set will include simulation results regarding the system performance of the LiDAR module (Module-4).	D2.3	Results	TXT, TSV, CSV, JPEG	<500 MB	Open	1. ICCS
Simulation studies of the photonic convolutional neural network (NN) modules	D2.4	Results	TXT, TSV, CSV, JPEG	<500 MB	Restricted/ Data accessible by the	1. ICCS

Data Description	Related Deliverable(s)	Type	Format	Estimated Volume	Access	IPR Owner
The data set will include simulation results regarding the system performance of the photonic convolutional NN modules (Module-5 and -6).					consortium members only	
Radiation patterns of single and multiple integrated grating coupler antennas. The data set will include simulation results of the expected performance of the grating coupler-based single- and antenna arrays.	D2.3, and D3.1	Results	.TXT, .JPG	<250 Mb	Restricted/ Data accessible by the consortium members only	7. Optagon
Evaluation of beamforming algorithms The data set will include simulation results from the performance evaluation of the beamforming algorithms that will apply to the optical network of LOLIPOP.	D2.3	Results	TXT, TSV, CSV, JPEG	<250 MB	Restricted/ Data accessible by the consortium members only	7. Optagon
3D mechanical modelling and thermal simulation of LOLIPOP modules The data set will include results of thermal simulation and step files of the 3D mechanical design	D2.5	Results	JPEG, .docx, PNG, STEP	<250 MB	Restricted/ Data accessible by the consortium members only	7. PHIX
WP3. Monolithic and heterogeneous integration on photonic platforms						
Characterization results of unpackaged single high-speed modulators on LNOI platform. The data set will include the response of the LNOI-based modulators.	D3.2	Results	TXT, TSV, CSV, JPEG	<250 MB	Restricted/ Data accessible by the consortium members only	12. CSEM
Characterization results of unpackaged periodically poled waveguides on LNOI platform. The data set will include the response of the PPLN components.	D3.2	Results	TXT, TSV, CSV, JPEG	<250 MB	Restricted/ Data accessible by the consortium members only	12. CSEM
Characterization results of unpackaged single PZT-based phase actuators. The data set will include the frequency response of PZT-based phase actuators.	D3.3	Results	TXT, TSV, CSV, JPEG	<250 MB	Restricted/ Data accessible by the consortium members only	2. LXI
Characterization results of low- and high-speed Ge PDs on TriPleX	D3.4	Results	TXT, XLSX, PPT	< 150 MB	Restricted/ Data accessible	5. UTwente

Data Description	Related Deliverable(s)	Type	Format	Estimated Volume	Access	IPR Owner
This data set includes the files generated by the characterization of the Ge-based PDs.					by the consortium members only	
WP4. Hybrid integration on TriPleX platform						
Assembly and packaging of LOLIPOP prototypes The data set will include step files of the mechanical designs and reports containing images of the completed prototypes.	D4.1, D4.2, and D4.4	-Results -LNOI and TriPleX chips, images	-TXT, TSV, CSV, JPEG - Optical images of the LNOI and TriPleX chips	<250 MB	Restricted/ Data accessible by the consortium members only	5.PHIX
Testing of the micro-transfer printing technology for LNOI-on-TriPleX platform The data set includes high level functional design, simulation and testing data of the μ-transfer printing method.	D4.3	Results	TXT, PPT, JPEG, GDS	<250 MB	Restricted/ Data accessible by the consortium members only	4.TYN
WP5. Front-end electronics and packaging						
High-speed PCBs for LOLIPOP prototypes The design of PCB motherboards with DC and RF access lines for the assembly of PICs and ICs and module packaging.	D5.1, and D5.3	Results and prototypes	GERBER, .DOC	<1 GB	Restricted/ Data accessible by the consortium members only	9. IMEC
Circuit schematics of the electronic integrated circuits This dataset contains the schematics of the driver and TIA chips including all transistor-level buildings blocks and their interconnections.	D5.1, and D5.3	Results and prototypes	GERBER, .DOC	< 1 GB	Restricted/ Data accessible by the consortium members only	9. IMEC
WP6. System integration and testing of LOLIPOP modules						
LOLIPOP control electronics This set of experimental data will contain the characterization measurements of the LOLIPOP control electronics	D6.1, and D6.4	Results	TXT, TSV, CSV, JPEG	<250 MB	Restricted/ Data accessible by the consortium members only	7. Optagon
Characterization and performance evaluation of LOLIPOP prototypes (Module-1 and Module-2) This data set will contain the experimental results from the system level evaluation of	D6.2, and D6.5	Results	TXT, TSV, CSV, JPEG	<500 MB	Restricted/ Data accessible by the consortium	8. Polytec

Data Description	Related Deliverable(s)	Type	Format	Estimated Volume	Access	IPR Owner
LOLIPOP prototypes Module-1 and Module-2.					members only	
Characterization and performance evaluation of LOLIPOP prototype (Module-3) This data set will contain the experimental results from the system level evaluation of LOLIPOP prototype Module-3.	D6.3	Results	TXT, TSV, CSV, JPEG	<500 MB	Restricted/ Data accessible by the consortium members only	**10. QuiX**
Characterization and performance evaluation of LOLIPOP prototype (Module-4) This data set will contain the experimental results from the system level evaluation of LOLIPOP prototype Module-4.	D6.6, and D6.8	Results	TXT, TSV, CSV, JPEG	<500 MB	Restricted/ Data accessible by the consortium members only	**12. CSEM**
Characterization and performance evaluation of LOLIPOP prototypes (Module-5 and Module-6) This data set will contain the experimental results from the system level evaluation of LOLIPOP prototypes Module-5 and Module-6.	D6.7, and D6.8	Results	TXT, TSV, CSV, JPEG	<500 MB	Restricted/ Data accessible by the consortium members only	**1. ICCS**

5 Allocation of resources

There are no immediate costs anticipated to make the datasets produced FAIR. The open datasets will be deposited in the Zenodo repository and will be preserved for at least 5 years after the conclusion of the project and in-line with the Deposit Data Policy of the European Commission. Any unforeseen costs related to open access to research data in Horizon Europe are eligible for reimbursement during the duration of the project under the conditions defined in the G.A. Article 6 and Article 6.2.C.3.

ICCS (Project Coordinator) is the person responsible for the original definition of DMP. All the consortium partners shall support the Data Management protocol by providing the types of generated data that have been described in the methodology of the DMP report.

Each LOLIPOP partner should respect the policies set out in this DMP. Datasets must be created, managed, and stored appropriately and in line with European Commission and local legislation. Dataset validation and registration of metadata and backing-up data for sharing through repositories is the responsibility of the partner that generates the data in the respective Work Package.

6 Data security

The data storage facilities of the members of LOLIPOP project are summarized in the table below.

Table 3. Description of data storage facilities of LOLIPOP beneficiaries

Beneficiary Short Name	Description of Data Storage Facility
1. ICCS	University-owned server at ICCS (PCRL). The server offers real-time data mirroring through RAID (redundant array of independent disks) and weekly backups to external disk drives. The Data will be preserved for at least 5 years after the project ends and the associated costs will be covered by ICCS through our own funds. Data will be also stored at the ICCS's SharePoint Documents Repository devoted to LOLIPOP by the project end.
2. LXI	The data storage is on a LioniX-owned server which has daily off-site backups. All original project data will be archived for at least 5 years after the project ends. Data which is non-confidential and relevant for the partners will also be stored on the project SharePoint.
3. PHIX	All original data will be stored on local hard drives (with regular backups) and archived on a local network such that these are available and accessible for all employees involved in the project. Processed copies of this data will be maintained on the PHIX servers and network drives. All data will be managed under the FAIR (Findable, Accessible, Interoperable and Reusable) principles during the project for the researchers in the team. Raw and processed files will be stored on local hard drives (with regular backups) and archived on a local network for at least 10 years and will be accessible to the LOLIPOP consortium. For the information that is used in publications, PHIX will upload the non-confidential original data onto the SharePoint (and MS TEAMS Workspace) or provide download links to other researchers upon request. Digital archives with sufficient storage capacity with adequate data transfer bandwidth will be made available during the length of the project and thereafter.
4. TYN	The data storage is on a TYN-owned server which has daily off-site backups. All original project data will be archived for at least 5 years after the project ends. Data which is non-confidential and relevant for the partners will also be stored on the project SharePoint.
5. UTwente	The data storage is on a UTwente-owned server which has daily off-site backups. All original project data will be archived for at least 5 years after the project ends. Data which is non-confidential and relevant for the partners will also be stored on the project SharePoint.
6. Superlum	The data storage is on a Superlum-owned server which has daily off-site backups. All original project data will be archived for at least 5 years after the project ends. Data which is non-confidential and relevant for the partners will also be stored on the project SharePoint.
7. Optagon	The data storage is on an Optagon-owned server which has daily off-site backups. All original project data will be archived for at least 5 years after the project ends. Data which is non-confidential and relevant for the partners will also be stored on the project SharePoint.
8. Polytec	The data storage is on a Polytec-owned server which has daily off-site backups. All original project data will be archived for at least 5 years after the project ends. Data which is non-confidential and relevant for the partners will also be stored on the project SharePoint.
9. imec	The data storage is on a imec-owned server which has daily off-site backups. All original project data will be archived for at least 5 years after the project ends. Data which is non-confidential and relevant for the partners will also be stored on the project SharePoint.
10. QuiX	The data storage is on a QuiX-owned server which has daily off-site backups. All original project data will be archived for at least 5 years after the project ends. Data which is non-confidential and relevant for the partners will also be stored on the project SharePoint.
11. IRIDA	The data storage is on an Irida-owned server which has daily off-site backups. All original project data will be archived for at least 5 years after the project ends. Data which is non-confidential and relevant for the partners will also be stored on the project SharePoint.
12. CSEM	The data storage is on a CSEM-owned server which has daily off-site backups. All original project data will be archived for at least 5 years after the project ends. Data which is non-confidential and relevant for the partners will also be stored on the project SharePoint.

7 Ethical aspects

LOLIPOP partners are to comply with the ethical principles as set out in the G.A. Article 14 which states that all activities must be carried out in compliance with:

a) Ethical principles (including the highest standards of research integrity — as set out, for instance, in the European Code of Conduct for Research Integrity (European Science Foundation, 2011) — and including avoiding fabrication, falsification, plagiarism or other research misconduct) and

b) Applicable international, EU and national law including the EU Charter of Fundamental Rights and the European Convention for the Protection of Human Rights and Fundamental Freedoms and its Supplementary Protocols.

The LOLIPOP consortium will conduct this action adhering to the fundamental principle of research integrity —as set out in the European Code of Conduct for Research Integrity[1].

This implies compliance with the following principles:

- **reliability** in ensuring the quality of research reflected in the design, the methodology, the analysis, and the use of resources

- **honesty** in developing, undertaking, reviewing, reporting, and communicating research in a transparent, fair and unbiased way

- **respect** for colleagues, research participants, society, ecosystems, cultural heritage, and the environment

- **accountability** for the research from idea to publication, for its management and organization, for training, supervision and mentoring, and for its wider impacts and means that beneficiaries must ensure that persons carrying out research tasks follow the good research practices including ensuring, where possible, openness, reproducibility and traceability and refrain from the research integrity violations described in the Code.

LOLIPOP project does not involve the use of human participants or personal data in the research and therefore there is no requirement for ethical review.

7.1 Confidentiality

LOLIPOP beneficiaries must and will retain any data, documents, or other material as confidential during the implementation of the project. Further details on confidentiality can be found in the G.A. Article 13 along with the obligation to protect results in Article 16.

8 Other issues

No other national/funder/sectorial/departmental procedures for data management are being used by the LOLIPOP partners.

---

[1] https://allea.org/code-of-conduct/

9 Conclusions

This document has provided the first version of the Data Management Plan defined by LOLIPOP consortium in the first nine months of the project. The DMP has identified several data sets which will be generated within the project and, for each of them, has defined the strategy for the documentation, open access sharing and maintenance of the associated data, according to the guidelines provided by the EC in the Horizon Europe Programme.

This document will be updated during the project lifecycle, to reflect any possible changes and additions in the data sets, as well as any refinement in the strategy to maximize the sharing and re-use of the project outcomes. These possible changes should also be reflected in the periodic reports on the LOLIPOP project. In particular, at least three further versions of the DMP will be released: one in M18, one in M30, and the final version at the end of the project in M42.

LOLIPOP Project: 101070441

List of Tables

Appendix I - Example Metadata File Template

This metadata file was generated on *<insert date>* by *<insert name>*

GENERAL INFORMATION

1. Title of Dataset:

2. Dataset Identifier in Repository:
3. Responsible Partner:
4. Author Information:
   Investigator Contact Information Name:
   Email:

   Supervisor Contact Information Name:
   Email:

   Co-Supervisor Contact Information Name:
   Email:

5. Date of data collection: ……………………………………
6. Geographic location of data collection (where was data collected?): ……………………………………
7. Title of project and Funding sources that supported the collection of the data: ………………

SHARING/ACCESS INFORMATION

1. Licenses/access restrictions placed on the data: ………………………………………………………………….
2. Link to data Repository: …………………………………………………………………………………………..
3. Links to other publicly accessible locations of the data: ………………………………………………….
4. Links to publications that cite or use the data: …………………………………………………………………
5. Was the data derived from another source? If yes, list source(s): …………………………………………

DATASET & FILE OVERVIEW

1. This dataset contains X sub-dataset as listed below:
   A. Datasheet name:
   B. Datasheet name:
   C. Datasheet name:
   D. Datasheet name:
2. What is the status of the documented data? – "complete", "in progress", or "planned"
3. Plans to update the data

METHODOLOGICAL INFORMATION

1. Description of methods used for experimental design and data collection: *<Include links or references to publications or other documentation containing experimental design or protocols used in data collection>*……………………………………………………………………………………

2. Methods for processing the data: *<describe how the submitted data were generated from the raw or collected data>*…………………………………………………………………………………..
3. Instruments and software used in data collection and processing-specific information needed to interpret the data: ………………………………………………………………………………………………

4. Standards and calibration information, if appropriate: …………………………………………………………….
5. Environmental/experimental conditions: …………………………………………………………………………………

Describe any quality-assurance procedures performed on the data: ……………………………………..