Deliverable Report

Deliverable No: D9.2

Deliverable Title: Data Management Plan

Grant Agreement number: 899544

Project acronym: PHOQUSING

Project title: PHOtonic Quantum SamplING machine

Project website address: www.phoqusing.eu

Name, title and organisation of the scientific representative of deliverable's lead beneficiary (task leader):
Prof. Fabio Sciarrino, Università di Roma "La Sapienza"
Email: fabio.sciarrino@uniroma1.it

Deliverable table

Deliverable no.	D9.2
Deliverable name	Data Management Plan
WP no.	9
Lead beneficiary	1 (UNIROMA1)
Type	ORDP: Open Research Pilot
Dissemination level	Public
Delivery date from Annex I	Month 6
Actual delivery date	28 February 2021

What was planned *(from Annex I:)*
D9.2: Data Management Plan [6]
A structured Data Management Plan will be prepared and will include the best standards for the generated data and assess their suitability for sharing and reuse in accordance with official guidelines.

1. Data Summary

The objective of the project is to develop two quantum sampling machines for hybrid quantum computing applications. Data will be divided in three main parts: (1) data generated for purposes of investigating the classes of algorithms which can be solved via the implemented hardware, (2) data for the development and usage of the sampling machines, and (3) data investigating the hybrid quantum computing.

Types of data will include: Software, reports (e.g. report on validation techniques for multiphoton interference, classification and development of sampling tasks, report system construction and integration, etc.), technical scheme and specifications for the two sampling machines, operator manual for the developed system. In addition, reports on financial analysis and end user analysis.

Formats will include:
Raw data: binary and numerical data.
Software and codes.
Reports in PDF and Office documents.

Reuse of data: Specifications of the components at the basis of the developed system will be re-used for system design.

Origin of data: All software and reports originate with the partners developing the system.

Expected size of data: ~ 100GB (very preliminary estimation, mainly from raw data to be exchanged)
Utility of data: Data on technical aspects will be useful internally within this project for development and characterization of the system.
General use will be for the long-term goal implementation of the sampling machines QOLOSSUS and QALCULUS.

2. FAIR data

2. 1. Making data findable, including provisions for metadata

Data will have unique identifier.
Publications will have appropriate search keywords.
Clear version numbers will be provided.

Metadata includes: Digital Object Identifiers (DOI) for published papers, preprint identifiers, cross-reference between repository preprint and published version, institutional repository identifiers.

2.2. Making data openly accessible

Three levels of accessibility for the data will be employed:

1) Data and software generated for the development of the sampling machines will be for internal use only. Specifications and manual will be for internal use. Data developed will be stored locally and shared between the involved partners.
2) Gold open access will be preferred for publications. If not possible, green open access will be granted for all publications through institutional repositories and/or arXiv repository (https://arxiv.org). Links redirecting to publications will be posted in the dedicated section of the project website.

3) Open access to data and software supporting publications, or additional material generated within the project, will be eventually made available through specialized portals (e.g. GitHub) upon agreement between the involved partners,

2.3. Making data interoperable

Compliance with currently available open software applications (e.g. Python, or other open-source programming languages) will be preferred for software development.
Data shared will be formatted by using standard methodology and formats.
Reports and publications will be prepared by using standard terminology and methodologies.

2.4. Increase data re-use (through clarifying licences)

Re-usage of specifications of the developed system will depend on the project outcomes.
Data produced for the development of the prototypes will be for internal use only.
Re-usage licenses for data referring to publications will be evaluated within the course of the project.
Re-usage of open source software made available via dedicated portals will be granted.

3. Allocation of resources

Costs for making data FAIR will be mainly due to the cost of gold open access publications.

Costs will be covered from the project budget.

Person responsible for data management: Each beneficiary will be responsible for their part of the data.
Resources for long term preservation (costs and potential value, who decides and how what data will be kept and for how long) are still being discussed.

4. Data security

Provisions for data security (including data recovery as well as secure storage and transfer of sensitive data) will be based on standard university procedures.

Safe storage of data in certified repositories for long term preservation and curation: Under consideration.

5. Ethical aspects

No ethical issues.