Deliverable D1.3 – Data Management Plan (GA No. 101021723)

Work Package(s)	WP 1
Task(s)	D1.3
Dissemination Level	Public
Due Date	31 October 2021
Actual Submission Date	29 October 2021
WP Leader	IDEAS Science Ltd.
Task Leader	DataSenseLabs Ltd.
Deliverable Leader	Anna Pálhalminé Mező (anna.mezo@datasenselabs.net)
Contact Person	Anna Pálhalminé Mező (anna.mezo@datasenselabs.net)
Project coordinator	Györgyi Bela (gyorgyi.bela@ideas-science.com)

Document History

Revisions	Author(s)	Date	Description
Version 1.0	Anna Pálhalminé Mező, János Pálhalmi	27 Oct 2021	Initial version of DMP
Version 2.0	Reviewed by Györgyi Bela	28 Oct 2021	Reviewed version

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No: 101021723

Table of Contents

Introduction

HoloZcan brings a new tool for security actors (police, relief workers, disaster managers, crisis managers, stakeholders responsible for public safety, critical infrastructure, and service providers) notably in the fields of autonomous detection and response capabilities.

The project will increase (environmental and exhaled) bio-aerosol sensing/measurement capability of CBRN practitioners by developing a high resolution, large throughput, automatic and highly portable detection system for making automatic classification of pathogens and particles.

HoloZcan develops of a novel holographic microscopy and imaging technology for rapid and cost-efficient screening of potential biological threats and unknown, potentially dangerous substances, combined with methods of artificial intelligence and machine learning. It establishes a framework of a dynamic feature selection and validation algorithm to support the continuous innovation capability of the system in the field of adaptive learning and database optimization for specific bioinformatic applications. The project also develops comprehensive and innovative means of respiratory, ventilation and environmental biological data sampling that can be used in real-time, standoff or in mobile bio-detection context.

The project indicates the HoloZcan technique versatility for a wide range of applications and demonstrates its technical feasibility. The project responses to the actual needs of European practitioners and technological gaps identified by the ENCIRCLE project as indicated in the ENCIRCLE Catalogue of Technologies and addresses several shortcomings of the current approaches to bio-threat agent detection.

The HoloZcan project applies a flexible adaptive approach to design and CBRN practitioners are engaged as project partners or as external stakeholders in the process.

The present document is a public deliverable of Work Package 1 (D1.3). It describes the complete lifecycle of all the data that are generated, used, and/or collected by the HoloZcan project: collection, management, preservation, archiving, sharing.

It contains information on:

- the handling of research data during & after the end of the project
- what data will be collected, processed and/or generated
- which methodology & standards will be applied
- whether data will be shared/made open access and
- how data will be curated & preserved (including after the end of the project)

Data Summary

The HoloZcan project will generate research data in the following Work Packages:

WP2: Biohazard classification system development

WP3: Bio Sampling and System Maintenance

WP4: Integration, adjustment, calibration and validation of the instrumentation

WP5: User engagement, dissemination, exploitation, and training

The data will be subject to change, like addition or removal of datasets, as needed by the project development. The data will be generated and stored in several different formats, which include: design files, text documents, image files, computer code, tables, surveys, etc.

Data that will be included in the Open Research Data Pilot (ORDP) will include:

- Open Access scientific journal materials,
- Conference materials,
- Workshop materials,

and certain datasets produced during the lifetime of the project - to be determined later.

Due to the nature of the project topic, part of the data to be generated by the HoloZcan project is going to be Confidential. This makes a strong limitation on what can be and will be shared in the Open Research Data Pilot (ORDP).

The Grant Agreement describes (29.3 Open access to research data):

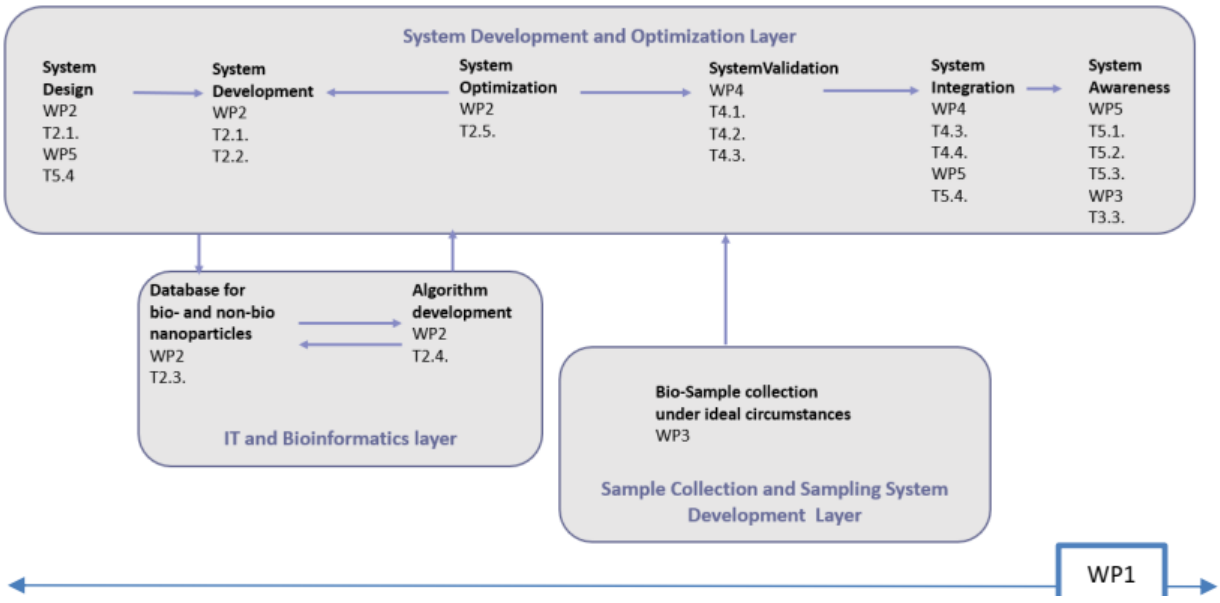
“As an exception, the beneficiaries do not have to ensure open access to specific parts of their research data under Point (a)(i) and (iii), if the achievement of the action's main objective (as described in Annex 1) would be jeopardised by making those specific parts of the research data openly accessible.” (GA Pg. 48)

The consortium members are committed to responsibly addressing the safety and security issues that arise and will not publish articles or databases where the potential harm of publication outweighs the potential social benefit. Before any publication is made, we conduct a security assessment, in which we evaluate the pros and cons.

The HoloZcan scientists are aware of their personal responsibilities to consider the balance of risks and benefits in research they consider undertaking. The project's security advisers help assess the safety risks and the social benefits expected from the experiments or dissemination activities.

The HoloZcan research does not provide knowledge, materials or technologies that can be used in the field of crime, cybercrime or terrorism. The results of the research cannot be used for biological weapons and the means to deliver them.

Fig.1. R&D DataFlow generation related to the work package structure: GA 3.1. Fig.6. page: 179



Planned Data Generation / Usage

WP2: Biohazard classification system development

Lead beneficiary: Sioux

Participants: IDEAS Science, DataSenseLabs, ZugMed, POLIMI, LODZ, KSP, DMI Associates, Institute Pasteur

Hardware-design related data – Sioux workflow:

Kind of data	**File formats**	**Storage**	**Modifications**	**Rentention duration**	**Shared with**
Experiments on test setup(s)	Images (.tif), configuration (.txt)	Local	Growing	Until design decision	Partners, on request
Algorithms, software	Source code (.m, .py)	Sioux internal GIT, Project GIT	Version-control	Long term	With partnes, via Project GIT
Hardware design	Siemens NX CAD models	Siemens versioning	Version-control	Long term	Partners, on request
QA / calibration of a single microscope	Tables, graphs (.pdf)	Shipped with product	Fixed	Until shipment	Receiver of the specific item.

The QA of the manufactured hardware is covered by the Sioux' Quality Management System (QMS).

The procedures are stored in the Device Master Record (DMR). Device-specific test reports will be made.

These test reports will be stored in the Device History Record (DHR). Both the DMR and DHR are stored in Siemens' Teamcenter.

Teamcenter has integrated version control which is used by Sioux.

Reused data:	Yes, databases and publicly shared software codes related to specific scientific publications.
Data Origin:	Spectrometer output for characterizing the light source.
Simulation:	Computer simulation outputs generated in the iterative workflows of software development, testing and algorithm development (WP2 and WP4).
Derived/compiled data:	Results of holographic image analysis, both from measurements and simulations.
Datasets are:	Growing: Data from experimental setups is added for each new setup. Revisable: Calibration data for e.g. the exact geometry of the setup
Quantity:	Current size: 500 MB, might grow when super-resolution, multiple wavelengths are added.
Local data storage:	Local copy on SSD for fast processing, on PCs with user identification, password and firewall/virus protection put in place.
Long term data value:	The experimental data is useful up to the point that hardware design decisions are made. The calibration results need to be stored, the raw data itself not.

Purpose – IP workflow:	Building of a reference library of images of biological particles (bacteria). Objectives: Reference library of images for AI training, system calibration and validation.
Data Type and Formats:	measurement results, tables (.xlsx, .csv), lab notes (.docx, .pdf, .txt), audiovisual – images (.tiff, .png)
Data Origin:	Observational: Serological status of COVID-19 from volunteers giving exhaled breath condensate for research purpose. Some of the donors will be patients and others healthy volunteers. Pseudonymized data., Experimental: Bioimage and physical object image databases generated in WP2 and WP4.
Dataset is:	growing, the library of images can be enriched over time for further algorithms development.

Quantity:	Microscope slides: the number will depend on the needs identified once the database building will start.
Data Storage and Security:	type of storage: Local and temporary storage in Institut Pasteur storing server, which is safe and confidential. HoloZcan data server. Institut Pasteur storage servers are limited to local individual users or groups of users and password protected. Based on the needs different safe backup policies are available: backup (snapshots on the same server to restore accidentally removed files) or backup plus replication (copy on a different server on the other side of the campus).

Purpose – LODZ Workflow	Database building for biological and physical nanoparticles
Objectives:	1. Pathogen database building 2. Contaminants database building
Type and Format of Data:	Text: Laboratory notes, research notes. .docx, .pdf, .txt Numeric: tables, .xls, .xlsx Audiovisual: images, .jpg, .tiff
Discipline specific information:	Biological data, in image formats: .jpg, .tiff
Reference or Canonical Links:	Reference database containing different types of microorganisms, noise background and physical objects
Dataset is:	Fixed: Datasets of biological and physical objects
Experimental:	Images generated in WP2, WP4
Quantity of data:	Depends on the exact format – Images of microscopic slides
Derived/Compiled:	Image analysis results
Data Storage and security:	Part of the data is stored locally on LODZ servers, with user and permission management set up. The password protected accounts are only available to the specific user. The University data storage and security practices apply. Part of the data will be stored on the central Consortium data server located in the central office of IDEAS Science.

Long term data value:	Confidential datasets and databases will be available only for the consortium partners for development purposes only. The non-confidential part of the datasets, databases and algorithms can be used for publishing open source scientific materials serving the scientific and tech. community in the field of microscopic R&D, bio-medical image analysis and algorithm development and also CBRN use-case developers working in the field of bio-threat detection.

WP3: Bio Sampling and System Maintenance

Lead Beneficiary: ZugMed

Participants: IDEAS Science, DataSenseLabs, POLIMI, LODZ, Sioux, KSP, DMI Associates, Institute Pasteur

Purpose – ZugMed workflow:	Sampling particles from a tube through a liquid conditioning to a microscope slide.
Objectives:	Collect samples from the air, Store sample in a liquid, transfer the samples from the liquid to a slide. The sampling data collection is in line with the objective of the project, as indicated in the WP3.
Type and Format of Data:	Text (ID_YYYYMMDD_GPS coordinates) on sampler stickers in .docx, .pdf, .txt label format, Numeric (table of potential pathogens on a sheet, table of environmental factors on a sheet) in .xlsx, .csv formats. Images in .jpg format of 3D design project of a prototype, which is a working concept model. Complete visuals with photos from every angle. Simulated 3D design models (prototype): 3D design plans are created for the working concept model and potential prototype for creation and assembly.
Discipline specific information:	economics, biological specific standard for respiratory monitoring medical technologies.
Instrument specific:	3D printer input files (CAD) – 3D printed plastic.
Quantity of data:	flexible, to be identified during the project
Observational data:	Data is going to be collected by first responders participating in the project: Warsaw police and Institut Pasteur
Experimental:	Provided by the participants, including the University of LODZ.
Simulation:	Simulated Data is always the first step set by CBRN engineers working on a concept model.
Derived/Compiled:	Test image compiled in test data can be considered as input for the sampling development of a prototype in a laboratory system.
Data Storage and security:	Offline storage, Our prototype design environment is currently working offline so there is no actual need for an online data security policy, however we plan to set a data security policy for the use of hard drives and the back-up system.

Long term data value:	Data will be useful for people using the DHM. The collaborative potential results of a limited pathogens database may be useful for long term comparisons of different pathogens.

WP4: Integration, adjustment, calibration and validation of the instrumentation

Lead Beneficiary: DataSenseLabs

Participants: IDEAS Science, ZugMed, POLIMI, LODZ, Sioux, KSP, DMI Associates, Institute Pasteur

Purpose – DataSenseLabs workflow:	System validation and calibration data. System fine-tuning for specific applications.
Objectives:	Dimensional calibration by scalable nano-, micro-object standards. Computer simulation dataset for software development and for software testing. Database of different microorganisms to support system validation and software development and testing. Non-Confidential database to share scientific findings
Data Types and Formats:	Text: .docx, .pdf, .txt, (Word notes, txt files, etc.), Numeric: .mat, .cs, .xlsx, .txt, computer code (tables, measurement results, measurement notes) Visual: .tiff, .png, HDF5 (measurement images, image data, simulation visualization results)
Simulated output:	Models - Equation in pseudocode, intensity values in relative (arbitrary) units and phase values in complex number output. Model type: Metadata within the computer code Computer code: .m, .py, Data Type: Numeric, image, database Format: .mat, .py, .csv., .txt, .tif., .png, HDF5
Discipline specific information:	Metrology and software development. Format: .mat, .py, HDF5, .tif, .png
Re-used data:	Yes, we plan to use databases and publicly shared software codes related to specific scientific publications
Experimental data origin:	Bioimage and physical object image databases generated in WP2 and WP4.
Simulation data origin:	Computer simulation outputs generated in the iterative workflows of software development, testing and algorithm development (WP2 and WP4).

Derived/compiled data:	Results of holographic image analysis, results of machine-, and deep-learning algorithms originated from the simulated, the bioimage and physical image databases. Results of the holographic image analysis, results of machine-, and deep-learning algorithms based on simulated holographic datasets.
Revisable/ Growing/ Fixed datasets:	Revisable: bioimage and background noise database to develop and train machine-, and deep-learning algorithms. Fixed: datasets of biological and physical objects used for calibration, testing and validation.
Quantity of data:	Approximately in the Terabyte range, it depends on the exact format of the databases to be specified during the project.
Long term data value:	Confidential datasets and databases will be available only for the consortium partners for development purposes only. The non-confidential part of the datasets, databases and algorithms can be used for publishing open-source scientific materials serving the scientific and tech. community in the field of microscopic R&D, bio-medical image analysis and algorithm development and also CBRN use-case developers working in the field of bio-threat detection. The non-confidential part of the datasets, databases and algorithms can be a strong foundation of further development to support biomedical and CBRN related R&D activity.
Local data storage:	On local high-performance computing unit, on SSD drives in software RAID 5 (capable of handling the loss of one full physical drive). Password protected local user accounts, with up-to-date antivirus, malware and firewall protection. Professional backup software used to create two more physical copies stored in different locations, offline, in a safe.

WP5: User engagement, dissemination, exploitation, and training

Lead beneficiary: DMI Associates

Participants: IDEAS Science, DataSenseLabs, ZugMed, POLIMI, LODZ, Sioux, KSP, Institute Pasteur

Purpose of the data – DMI and IDEAS Science workflow:	The process of designing adapted bio-detectors relies on a series of regular technical input from future users, with a special focus on "Stakeholders" that will voluntarily contribute by exchange throughout the entire project's duration with the Partners. The operation of collecting User's requirements and suggestions, takes multiple forms with a series of events and consultations and surveys organized.
Types of Data:	Each contributing User or Stakeholder will be identified with basic personal identification data. Following the presentation of technical details on the HoloZcan system, the participants to special focus group will have the opportunity to offer their comments and recommendations. Notes will be taken by the project or directly drafted by the Users. The data will bear a narrative dimension and consist mainly in observational information. Online survey results will also be collected in order to diversify the sources and variety of suggestions. All will be in text mode, with some identified to a specifical author and some will be only statistical or anonymous.
Purpose of the data:	The identification of user needs and requirements will allow to analyze by distinguishing and sorting the information, in order to regroup data in families of suggestions, allowing to then develop exactly the needed sub-categories of bio-detection devices as per future market expectations.
Project data storage and protection	All data collected, either electronically or on paper (and then scanned into PDF files), will be sent for a single copy version storage on Project HoloZcan's customized NextCloud server based in Hungary in a secured location. Project HoloZcan decided to use its services because it is designed to protect user data through multiple layers of protection. They have special webpage informing their customers about the protection levels offered: https://nextcloud.com/secure/

	Nextcloud aligns with industry standards such as Clause 14 of ISO/IEC27001-2013.
Deletion of data	The Users' and Stakeholders' input collected type of data will be deleted from the project server at project completion. No data will be retained in any possible format.

Compliance of DMP with Bio-Imaging Data Format Standards and Guidelines

The image output of quantitative microscopic methods has been used to describe and classify the patterns and features of different specimens in the field of biomedical and life sciences. Microscopic imaging technology is extremely evolving towards integrated hardware and software applications, which accelerates further innovations and generates challenges as well at several levels of interdisciplinary science and technology. One of the major challenges is the design and the management of data formats which can satisfy the requirements from different coexisting fields like scientific research, prototyping, industrial level development and generation of publicly available information to strengthen communication and knowledge transfer.

Several platforms (Global BioImaging, etc.) have been established to help and educate experts in the design and management of image data repositories coupled with metadata. Because of the wide range of quantitative imaging applications there are many different and application specific guidelines depending on the research domain and the level of prototyping and development process.

Since the HoloZcan project is the series of fundamentally interdisciplinary R&D workflows interconnected at several layers of the development, we chose to follow those generalized guidelines and best practices which can be implemented into the specific interdependent domains of the work packages, not forgetting the confidential nature of the development process. On the other hand, supporting the scientific and technology driven developer communities in the field of biomedical imaging is also an intention within the HoloZcan project, so we think it is crucial to build up a process for continuous monitoring of decision making related to the type and purpose of the data generated during the development. During this process the concepts of FAIR data principles: findability, availability, interoperability, reusability will be harmonized at the CBRN security level of the HoloZcan project (explanatory flowchart: Fig.2. page 18.).

Based on the above-mentioned harmonization and optimization concepts of openness and industrial level of security, the bioimage data ecosystem will be developed according to the best practices and experiences derived from the usability of the publicly available cell and tissue specific databases.

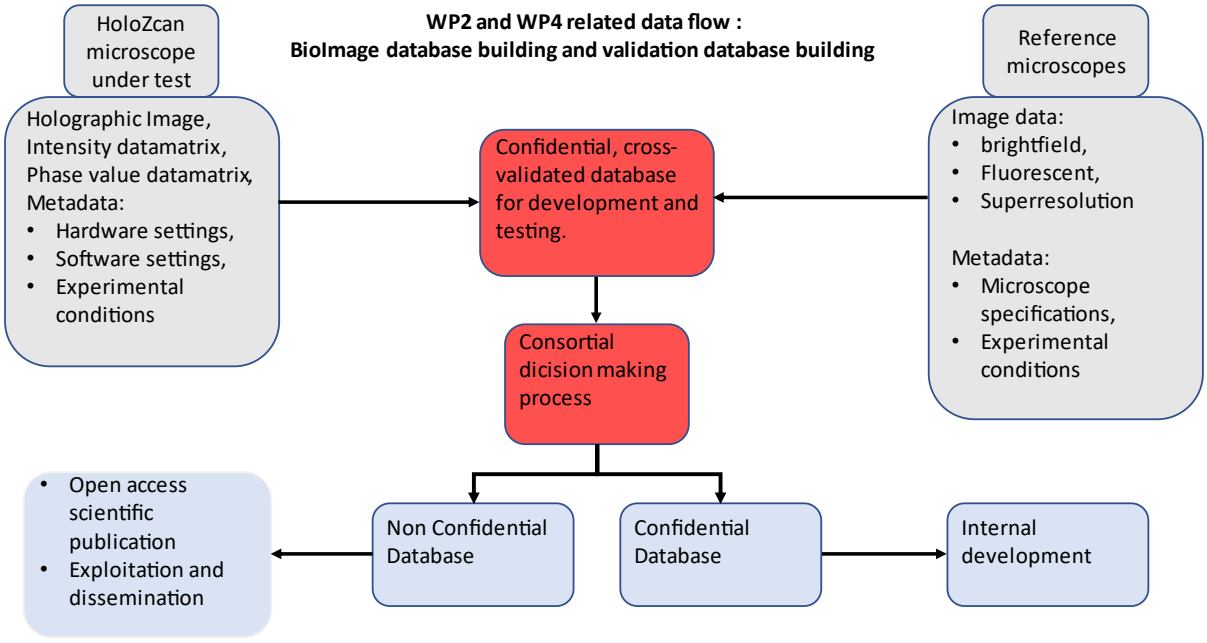
FAIR Data

WP2 - WP4 related FAIR data

Making data findable:

Due to the confidential nature of the HoloZcan project, the Consortium will decide upon which datasets will be subject to the FAIRification process. Datasets declared “confidential” or “classified” will not be made readily available to the public or third parties. In the case of confidential data proper licensing and IPR protection should be put in place, while in the case of classified data there is a high security risk involved with sharing that would compromise the HoloZcan final deliverable.

Fig.2. Structure of the data collection and decision making:



Regarding the decision making related to the open access FAIRification a process will be worked out involving the SAB (security advisory board) and the project Security Expert to integrate the responsibilities and the field of expertise.

Data will be identifiable via DOI – permanent link identifier.

Naming conventions:

Workflow name in short_holozcan_wp number_date_version.file extension

Search keywords approach:

Keywords are determined based on the database shared confidentially within the HoloZcan consortium: nomenclature of the topics under HoloZcan Zotero platform. Keywords are also determined based on the WP specific task and workflow descriptions (GA).

Versioning:

All final (major) versions of results will be stored in backup – traceability is ensured by timestamps and file version **naming practices**:

1. WP4 specific calibration, testing and validation results:

ctb_holozcan_wp4_date_verX_Y.extension

2. WP2 and WP4 specific software development related computer simulation results:

sim_holozcan_wp2_4_date_verX_Y.extension

3. WP2 and WP4 specific software development related code versioning:

sw_referring workflow name (for ie.: image processing, image reconstruction, machine-deep learning algorithm)_holozcan_wp2_4_date_verX_Y.extension

Procedures for metadata creation:

Content of metadata:

- Prototype under test: HoloZcan DHM microscope prototype versioning, circumstances of the measurement including the hardware and software related settings,
- Device used as a reference: referring data to the validation microscope used during the WP4 workflow, circumstances of the measurement and laboratory conditions, nomenclature of the biological and physical objects used for testing and database building, microscope settings, parameter values.

Standards will comply with the laboratory best practice guidelines and protocols.

Findable data:

The bibliographic metadata will be in a standard format and will include all of the following:

- the terms “European Union (EU)” and “Horizon 2020”;
- the name of the action, acronym and grant number;
- the publication date, and length of embargo period if applicable, and
- a persistent identifier

Naming conventions:

Workflow name in short_holozcan_wp number_date_version.file extension

Making data openly accessible:

The produced datasets that will be identified as shareable are planned to be placed in the Zenodo platform of OpenAIRE (Open Access Infrastructure for Research in Europe), EUDAT Collaborative Data Infrastructure, GitHub, and on the HoloZcan project website. The deposit will contain the data, including associated metadata, and other data. Information will be provided about tools and instruments at the disposal of the beneficiaries and necessary for validating the results, and also, if it is possible, the tools and instruments themselves will also be provided.

Joint ownership is planned among the consortium members according to the GA and CA regarding the confidential data (confidential results, codes and databases).

The sharing and publishing policy of the non-confidential part of the data will be monitored and decided upon by the Consortium members with the help of the Security Advisory Board.

Data kept closed:

According to the referring Data Security & Storage protocols and to the exploitation and dissemination plans.

Making data interoperable:

For indexing purposes keywords will be used. Keywords will be determined based on the database shared confidentially within the HoloZcan consortium: nomenclature of the topics under HoloZcan Zotero platform. Keywords are also determined based on the WP specific task and workflow descriptions (GA)

Making data interoperable:

The use of REMBI (Recommended Metadata for Biological Images) is planned – to be discussed during the project.

Data Licensing: Creative Commons Licenses, to be decided upon in a later project phase.

Increase data re-use:

Data usability for Third Parties: Data considered non-confidential will be made available via scientific open access portals (Zenodo, EUDAT, GitHub).

Confidential data will not be made available.

Quality Assurance (QA):

Continuous monitoring of the workflows related to WP2 and WP4, continuous monitoring of decision making related to data security and confidentiality. Controlled and documented by regular reports on QA meetings with the Coordinator (IDEAS Science).

There is no current time limit on the reusability of data.

Allocation of resources:

There is no separately allocated budget for the FAIRification process, however, there is allocated budget for certain partners and the Consortium leader for scientific publication.

Lodz and IDEAS Science can finance the cost of scientific publication in open-access peer-reviewed journals.

Data Security:

Local availability only, user and permission management in place. Backup process implemented to local drives/servers. Continuous monitoring of decision making is suggested to define which part of the datasets and databases are confidential or non-confidential. The decision-making process is defined within the general agreement of the consortium.

Certain data sets will be stored on the central project NextCloud server operated by the Consortium leader, IDEAS Science. These data will be available only to specified users via secure channels (sftp, etc.). The data will be situated on an encrypted drive in the central office of IDEAS Science.

WP3 related FAIR data

Data Openly available: Before any publication (inc data) is made, we conduct a security assessment, in which we evaluate the pros and cons. The HoloZcan scientists are aware of their personal responsibilities to consider the balance of risks and benefits in research they consider undertaking. The project's security advisers help assess the safety risks and the social benefits expected from the experiments or dissemination activities.

Data kept closed: also true for Institut Pasteur and LODZ university: Highly sensitive CBRN data is strictly confidential even if test data/inactivated pathogens.

Data will only be available for higher level institutions. Sampling data to be provided to the imaging / pathogen processing entity of the HoloZcan project.

Methods for data access: No active components, no executables, only static data, also for the sake of security. Simple file-based system stored off-line, Numeric/text. Only manual access is provided, access control manually recorded and controlled.

Data Interoperability: One of the key aims of the sampling strategy is creating a sample that can be transferred to the imaging system in order to facilitate the interoperability between the sampler and the imaging system: data transfer of sample to imaging system

Data usage timeline: Data can only be used during the timeframe that is the lifetime of the pathogen studied. Strictly confidential CBRN data for strictly controlled legal entities only. After the end of the project the test pathogens are to be destroyed. The data results are to be stored in the project repository for maintenance.

3rd party data usage: Data is confidential because it is part of a CBRN project. It can only be transferred to the competent superior entity.

QA: A key development factor of the sampling subsystem is the quality of the samples. A comprehensive advisory framework supports the development, by the University of Lodz and

Institute Pasteur. The sampling quality is measured by efficiency, availability, conservation, durability, and repeatability. Even though the HoloZcan project is a CBRN framework, still, we plan to follow ISO 9001 and potentially even ISO 13485 quality management standards to prepare for a potential spin-off on the long term in medical technologies as well.

Costs for making data FAIR: For the data to be eventually available to the public, financing may be needed through private equity.

WP5 related FAIR data:

Data Openly Available: Joint ownership among the consortium members according to the GA and CA regarding the confidential data (confidential results, codes and databases).

The sharing and publishing policy of the non-confidential part of the data will be monitored and decided upon according to the Data Security & Storage protocols.

Data kept closed: According to the Data Security & Storage protocols and to the exploitation and dissemination plans.

Methods for Data Access: write methods and tools:

methods and software tools that support the training activity will be made available to those who participate in the trainings or are interested in developing further training materials

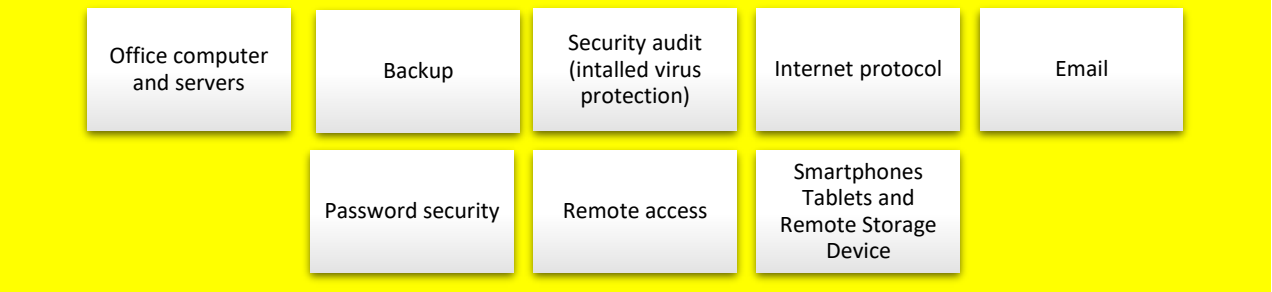
Repository for data deposit, metadata, documentation and code: NextCloud: https protocol / secured channel, authentication protected, Git repository for source code sharing: https protocol / secured channel, authentication protected, SFTP server for storing and sharing measurements: https protocol / secured channel, authentication protected

Access restrictions: Only partners will have access to the data. All repository and data sharing services are behind firewall (non-direct public IP), only the strictly required ports will be open

Backup: grade1 mirror and external backup

Additional Information

General Data Security Policy:



1. Office computers and servers: office computers have password protection + server data partitions are encrypted

2. Backup: using fault tolerant RAID systems and regular offline backups on external hard-disk

3. Security audit: installed virus protection, on Linux systems ClamAV is installed, on Windows computers we use various software

4. Internet protocol: All restricted data is sent over SSL

5. Email (server for data storage does not include mail server): we don't share sensitive information via e-mail, we use the NextCloud collaboration system located in the IDEAS Science main office, sftp and git over ssl to share documents, data and code with each other, and we put links to these resources into the e-mail body

6. Password security: Passwords must contain at least 8 characters, at least one number, and at least one non-alphanumeric character

7. Remote access: We don't use VPN, only HTTPS and SFTP protocols.

8. Smartphones Tablets and Remote Storage Devices: The storage of sensitive data on mobile devices is only temporarily allowed. After the current use, the relevant data must be deleted.

Intellectual Property Rights (IPR):

All data and information are subject to evaluation by the Consortium in connection with IPR protection, potential commercialization and security issues. The process is supervised by the Security Advisory Board and the project Security Expert, Marcin Niemcewicz.

The SMEs of the consortium (DataSenseLabs, IDEAS Science and DMI) are currently participating in the IP Scan process offered by European Union for Horizon 2020 projects. DataSenseLabs has already completed the process, while IDEAS Science and DMI are still working with their assigned legal experts regarding IPR protection and potential IPR issues. Therefore IPR questions will be addressed in an upcoming, amended version of the Data Management Plan.

Ethical Aspects

The Data Management Plan can be read in conjunction with the following ethics deliverables:

D6.1 -- H - Requirement No. 1 --submitted 31/07/2021

D6.3 -- HCT - Requirement No. 3 – submitted 31/07/2021

D6.4 -- POPD - Requirement No. 4 --submitted 31/07/2021

D6.5 -- POPD - Requirement No. 5 --submitted 31/10/2021

D6.6 -- EPQ - Requirement No. 6 --submitted 31/10/2021

D6.7 -- M - Requirement No. 7 --submitted 31/10/2021

References

A Global View of Standards for Open Image Data Formats and Repositories. Swedlow, Jason R., Pasi Kankaanpää, Ugis Sarkans, Wojtek Goscinski, Graham Galloway, Leonel Malacrida, Ryan P. Sullivan, et al. *Nature Methods*, 4 May 2021, 1–7. https://doi.org/10.1038/s41592-021-01113-7.

REMBI: Recommended Metadata for Biological Images—Enabling Reuse of Microscopy Data in Biology. Sarkans, Ugis, Wah Chiu, Lucy Collinson, Michele C. Darrow, Jan Ellenberg, David Grunwald, Jean-Karim Hériché, et al. *Nature Methods*, 21 May 2021, 1–5. https://doi.org/10.1038/s41592-021-01166-8.

Phase Contrast Time-Lapse Microscopy Datasets with Automated and Manual Cell Tracking Annotations. Ker, Dai Fei Elmer, Sungeun Eom, Sho Sanami, Ryoma Bise, Corinne Pascale, Zhaozheng Yin, Seung-il Huh, et al. *Scientific Data* 5, no. 1 (13 November 2018): 180237. https://doi.org/10.1038/sdata.2018.237.

Global BioImaging. Accessed 20 October 2021. https://globalbioimaging.org/.

BioImage Archive - EMBL-EBI' Accessed 13 September 2021. https://www.ebi.ac.uk/bioimage-archive/.

The FAIR Data Principles. Hagstrom, Stephanie. FORCE11, 3 September 2014. https://www.force11.org/group/fairgroup/fairprinciples.

FAIR, Open, and Free Does Not Mean No Restrictions. Jeffery, Keith G. *Patterns* 2, no. 9 (10 September 2021). https://doi.org/10.1016/j.patter.2021.100339.

Security Sensitivity Assessment

Objective

This form is related to the Security Sensitivity Assessment procedure which will assure that no sensitive information will be included in the publications and deliverables of the HoloZcan project.

Security sensitive information means here all information in whatever form or mode of transmission that is classified by Council Decision on the security rules for protecting EU classified information (2011/292/EU) and all relevant national laws and regulations. The information can be already classified, or such that it should be classified.

In practice the following criteria is used:

- Information is already classified
- Information may describe shortcomings of existing safety, security or operating systems
- Information is such, that it might be misused.
- Information that can cause harm to
    - European Union
    - a Member State
    - society
    - industry and companies
    - third country

citizen or an individual person of a country.

Document Information

Project:	HoloZcan: Deep Learning Powered Holographic Microscopy for Biothreat Detection on Field Grant Agreement No: 101021723
Deliverable:	D 1.3
Dissemination Level:	CO - Confidential, only for members of the consortium (including the Commission Services)
EU Project Officer:	RISCHITOR Patricia Elena
Actual Submission Date:	31/10/2021
WP Leader:	IDEAS Science
Authors:	Anna Pálhalminé Mező, Dr. János Pálhalmi

Assessment form for the main author

Please fill in the form below:

This is: pre-assessment □ final assessment ☒

List the input material used in the publication/deliverable: ---

List the results developed and presented in the publication/deliverable:

Clarification on how the consortium will address the EPQ--Requirements

The draft publication

☒ is attached to this statement

□can be found in link: --

This publication does include any data or information that could be interpreted as security sensitive. □ Yes ☒ No □ Not sure

If not sure, please specify what are the material / results that you are not sure if they are security sensitive? Why?

Date: 27 October 2021

Signature of the Responsible Author (Anna Pálhalminé Mező): ……

Comments from the SAB member

☒ The publication can be published as it is.

Comments:

No

□ Before publication the following modifications are needed: - -

Comments:

Date 27-10-2021

Name: On behalf of the Security Advisory Board (SAB) Dr. Marcin Niemcewicz

Signature of the member of the SAB ……