OSIRIS - dual Oct-baSed navIgation in Robot-assisted microSurgery
FWO DMP (Flemish Standard DMP)

1. Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data.

Dataset Name	Description	New or reused	Digital or Physical	Only for digital data Digital Data Type	Only for digital data Digital Data format	Only for digital data Digital data volume (MB/GB/TB)	Only for physical data Physical volume
Requirements List	Spreadsheet with the requirements for the development of an image-guided navigation system for robot-assisted microsurgery	Generate new data	Digital	Compiled data	.xlsx	< 100 MB	
Motion Generation Platform Hardware	List of materials, CAD Models and schematics for the development of a motion generation platform	Generate new data	Digital	Compiled data	.xlsx .par .asm .stl .kicad	<100 MB	
Motion Generation Platform Software	Source code for motion generation platform control	Generate new data	Digital	Software	.py .cpp	< 100 MB	
Intrument-Mounted Optical Fiber	Report methodology and list of materials for the development of an instrument-mounted optical fiber (fiber-OCT)	Generate new data	Digital	Compiled data	.xlsx	< 100 MB	
miOCT Raw Videos	OCT videos for the development of OCT super-resolution methods	Generate new data	Digital	Experimental	.jpeg .rosbag	< 100 GB	
Super-resolution miOCT Code	Source code for processing miOCT data to create super-resolution images	Generate new data	Digital	Software	.py	< 100 MB	
dual-OCT Calibration Videos	OCT videos from miOCT and fiber-OCT used to calibrate the two modalities	Generate new data	Digital	Experimental	.jpeg .rosbag	< 100 GB	
dual-OCT Calibration Code	Source code for dual-OCT calibration	Generate new data	Digital	Software	.cpp .py	< 100 MB	
dual-OCT Anatomy Videos	OCT videos from miOCT and fiber-OCT of anatomical structures	Generate new data	Digital	Experimental	.jpeg .rosbag	< 100 GB	
Anatomy CT-Scans	CT or micro-CT Scans of the ear or eye anatomy that was imaged with OCT	Generate new data	Digital	Experimental	.tiff .jpg .dcm	< 5 TB	
Anatomy CT-Scans from Open Datasets	CT or micro-CT Scans of the ear anatomy	Reuse data	Digital	Experimental	.tiff .jpg .dcm .nrrd	< 5 TB	

fiber-OCT Anatomy Landmark Tracking Code	Source code for anatomy tracking and reconstruction from fiber-OCT videos	Generate new data	Digital	Software	.cpp .py	< 100 MB	
miOCT Anatomy Landmark Tracking Code	Source code for anatomy tracking and reconstruction from miOCT videos	Generate new data	Digital	Software	.cpp .py	< 100 MB	
dual-OCT Motion Compensation Videos	OCT videos from miOCT and fiber-OCT of moving anatomies	Generate new data	Digital	Experimental	.jpeg .rosbag	< 100 GB	
dual-OCT Motion Compensation Sensors Data	Encoder data from robot on moving anatomies	Generate new data	Digital	Observational	.rosbag	< 100 MB	
dual-OCT Motion Compensation Code	Source code for robot control on moving anatomies	Generate new data	Digital	Software	.cpp .py	< 100 MB	

If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL etc.) per dataset or data type:

Anatomy CT-Scans from Open Datasets: https://doi.org/10.5281/zenodo.1473724
The 'OpenEar' dataset contains 8 temporal bone CT-scans registered with microslicing.

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? Describe these issues in the comment section. Please refer to specific datasets or data types when appropriate.

- Yes, human subject data

In the case of the Anatomy CT-Scans dataset, ethical approval shall be required for usage and processing of the data. However, the surgeons with whom we are collaborating will be responsible for acquiring these data or assisting with creating protocols for reusing data, considering the associated ethical issues and guidelines provided by the ethical committee.

Will you process personal data? If so, briefly describe the kind of personal data you will use in the comment section. Please refer to specific datasets or data types when appropriate.

- No

Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, …)? If so, please comment per dataset or data type where appropriate.

- Yes

The datasets have no immediate potential for commercial valorization, but may offer additional insight which could lead new IP. If human studies are conducted those will be research-initiated studies. This while contributing to the overall know-how these study results cannot be directly used e.g. for certification of the equipment.

- **dual-OCT Calibration and Motion Compensation Codes**: Integration of the dual-OCT system into robotic systems that require micro precision to perform tasks on moving structures.

- **Anatomy Landmark Tracking Codes**: Integration of anatomy landmark detection in OCT-based systems for surgical

navigation or patient diagnosis.

- **Motion Generation Platform Hardware and Software**: Simulation of millimetric or micrometric periodic motion for robot task performance validation on moving structures.

- **Instrument-Mounted Optical Fiber**: Development of medical graded instrument-mounted optical fibers for measuring distances and observing micro and millimetric structures from OCT imaging.

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material/Data transfer agreements/ research collaboration agreements)? If so, please explain in the comment section to what data they relate and what restrictions are in place.

- Yes

The OpenEar dataset (https://doi.org/10.5281/zenodo.1473724) is protected with the Creative Commons 4.0 license. Thus, usage requires proper referencing.

Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain in the comment section to what data they relate and which restrictions will be asserted.

- Yes

The OpenEar dataset (https://doi.org/10.5281/zenodo.1473724) is protected with the Creative Commons 4.0 license. Thus, usage requires proper referencing.

2. Documentation and Metadata

Clearly describe what approach will be followed to capture the accompanying information necessary to keep data understandable and usable, for yourself and others, now and in the future (e.g., in terms of documentation levels and types required, procedures used, Electronic Lab Notebooks, README.txt files, Codebook.tsv etc. where this information is recorded).

- **Image or video datasets**: a README.txt file will be written. This file will contain a complete description of the image acquisition methodology and image metadata (resolution, dimensions, and range scale). In case of DICOM images (.dcm), the DICOM file header will also contain such description.

- **Source code**: a README.txt file will be written. The source code (.cpp and .py files) will also be commented for clear understanding. This file will contain an extensive description of the code objective, compilation, and usage.

- **Experiments Materials and Methods**: a spreadsheet (.xsxl file) containing the description of the whole methodology.

Will a metadata standard be used to make it easier to find and reuse the data? If so, please specify (where appropriate per dataset or data type) which metadata standard will be used. If not, please specify (where appropriate per dataset or data type) which metadata will be created to make the data easier to find and reuse.

- No

There is no specific metadata for the datasets created in this project, thus, the documentation and metadata will follow a customized standard. Enough documentation will be written in order to ensure the reproducibility of experiments and results. The metadata will be inserted into the files described in the previous question.

3. Data storage & back-up during the research project

Where will the data be stored?

The image and video datasets will be stored in a secured storage on the archive (L-drive) by KU Leuven. Moreover, source code will be stored in KU Leuven's GitLab repositories.

How will the data be backed up?

Regular backups will be performed by the KU Leuven ICTS staff.

Is there currently sufficient storage & backup capacity during the project? If yes, specify concisely.
If no or insufficient storage or backup capacities are available, then explain how this will be taken care of.

- Yes

The storage offered by KU Leuven is a paid service and the research group has sufficient funds to pay for the storage service during the research and up until 5 years after the end of the research.
KU Leuven's GitLab also has sufficient storage capabilities for multiple repositories containing the source code. Considering each source code dataset requires one repository, 6 repositories will be needed, which is supported by GitLab.

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

All data and source codes will be managed in compliance with the IT guidelines, including password protection for all computers and external drives. The KU Leuven ICTS staff is also responsible for maintaining the security of the storage services. ICT maintains a dedicated folder structure offering only access if linked to the project. Therefore, only the researchers involved in this project will have access to the data and have to follow a two-factor authentication to access the GitLab repositories.

What are the expected costs for data storage and backup during the research project? How will these costs be covered?

There are expected costs for the usage of the secured storage offered by KU Leuven. A cost in the order of a few 100 Euros is expected. This cost can be easily covered by the RAS research group.

4. Data preservation after the end of the research project

Which data will be retained for at least five years (or longer, in agreement with other retention policies that are applicable) after the end of the project? In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...).

All the generated data will be preserved for at least 5 years after the end of the project.

Where will these data be archived (stored and curated for the long-term)?

The data will be archived in the KU Leuven secured storage service (L-drive) and KU Leuven GitLab Repositories for the long-term.

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?

The expected costs (few 100s or euro's) will be covered by the RAS research group.

5. Data sharing and reuse

Will the data (or part of the data) be made available for reuse after/during the project? In the comment section please explain per dataset or data type which data will be made available.

- Yes, in a restricted access repository (after approval, institutional access only, …)
- No (closed access)

After publication of the results, the OCT images and source codes will be available in a restricted access repository.
The anatomy CT-Scans will not be made available for reuse.

If access is restricted, please specify who will be able to access the data and under what conditions.

The restricted data will be accessible to the RAS members and can be made available under request by other research groups.

Are there any factors that restrict or prevent the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)? Please explain in the comment section per dataset or data type where appropriate.

- Yes, Intellectual Property Rights

The data will shed some insight in the technology that is developed. The potential to translate that technology will be explored in collaboration with big players in the field. Clear ownership of the data can help in discussions with such players.

Where will the data be made available? If already known, please provide a repository per dataset or data type.

Most likely through Mango a tool from the KUL's data-centrum.

When will the data be made available?

The data will be made available upon request and approval by the RAS Principal Investigator (Prof. Emmanuel Vander Poorten).

Which data usage licenses are you going to provide? If none, please explain why.

Source code will be under MIT License.

Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, you have the option to provide it in the comment section.

- No

What are the expected costs for data sharing? How will these costs be covered?

There are no expected costs for data sharing.

6. Responsibilities

Who will manage data documentation and metadata during the research project?

The researcher responsible by the project: Daniel Costa.

Who will manage data storage and backup during the research project?

The researcher responsible by the project: Daniel Costa.

Who will manage data preservation and sharing?

The RAS data manager: Dr. Mouloud Ourak.

Who will update and implement this DMP?

The researcher responsible by the project: Daniel Costa.