## DMP title

**Project Name** ICARUS - DMP title
**Project Identifier** 3E210520
**Grant Title** 3E210520
**Principal Investigator / Researcher** Hans Dierckx
**Project Data Contact** h.dierckx@kuleuven.be
**Description** Cardiac arrhythmias are a leading cause of death worldwide. Although different heart rhythm disorders can be recognized using the electrocardiogram, the precise three-dimensional spatiotemporal organization of cardiac activation patterns inside the heart during certain arrhythmia types is not well understood. Commonly, invasive catheter recordings are needed to record electrograms that hint at the location of abnormal sources or re-entrant activity. A pilot study in 2018 (Christoph et al., Nature) has shown that the electrical vortices in the heart driving the rhythm disorders can be reconstructed from mathematical analysis of the mechanical deformation in ultrasound (US) images. This proposal gathers experts at KU Leuven in ultrasound imaging, cardiac modelling and cardiology to bring this technology to clinical applications. Hereto, we will (a) lift US acquisition methods using state-of-art technology available at KU Leuven;(b) develop and perform inverse electromechanical modelling to reconstruct electromechanical activation sequences in the heart; (c) use this technology on patients to create personalized computer models (3D+time) and use the in silico reconstructions of spatiotemporal cardiac activation to optimize treatment (e.g. ablation, resynchronisation, pharmaceutics).
**Institution** KU Leuven

### 1. General Information

**Name of the project lead (PI)**
Hans Dierckx

**Internal Funds Project number & title**
Inverse modeling of Cardiac ARrhythmia sources based on UltraSound Imaging (ICARUS)
Project n°3E210520

### 2. Data description

**2.1. Will you generate/collect new data and/or make use of existing data?**

- Generate new data
- Reuse existing data

**2.2. What data will you collect, generate or reuse? Describe the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a numbered list or table and per objective of the project.**

| Type of data | Origin | Format | Volume |
|---|---|---|---|
| Synthetic datasets obtained by numerical simulation | Generated within the project by KU Leuven: relevant electrical model variables throughout the simulation | .var, .npy- | 1 GB x 100 = 100 GB |
| Output data (defined on mesh) from electromechanical modelling | Generated within the project by KU Leuven: estimated variables after optimization/inversion or data generated by forward models | .dat, .var | 5 MB x 100 = 500 MB<br>500 MB x 100 = 50 GB |
| Software for electromechanical modelling optimization/inversion and forward models | Generated within the project at KU Leuven: code for optimization/inversion and forward models | .py, .cpp | < 1 GB |

| | | | |
|---|---|---|---|
| 2D/3D Conventional and High Frame Rate Experimental Ultrasound Recordings | Generated within the project by KU Leuven: images/movies of the human hearts (patients and healthy volunteers), as well as phantoms and ex vivo tissue | Processed data (.dcm, .png, .avi, .xls)<br>Digital audio data (primary, raw) | < 3 TB |
| Diffusion Tensor Images (MRI images) | Generated within the project by KU Leuven: images of an ex vivo heart for estimation of the fiber orientation | .dcm | < 50 GB (estimated) |
| Intracardiac maps of arrhythmia patients | Generated within the project by KU Leuven: from arrhythmia patients undergoing intracardiac mapping for cathether ablation | TBD | < 20 GB (estimated) |
| Simulated Ultrasound Recordings | Generated within the project by KU Leuven | Digital audio data (primary, raw) | 20 GB x 100 = 200 GB |
| Electrocardiogram recordings | Generated within the project by KU Leuven | TBD | < 20 GB (estimated) |
| Ultrasound image acquisition, analysis and processing software | Generated within the project by KU Leuven: 2D/3D ultrasound beamforming sequences, myocardial strain and stiffness estimators, and estimator of cardiac fiber orientation. | .m files, C & C++ codes | <1 GB |
| Processed data from ultrasound recordings | Generated within the project by KU Leuven: strain and strain curves, activation maps, fiber orientation maps, stiffness estimates, etc. | Processed data (.mat, .xls, .jpeg, .avi, .gif) | ~100 GB |
| Lab books | Electronic lab notebook and paper notes | .pdf, .docx | <1 GB |
| Manuscripts, figures, reports, procedures, and presentation materials | Generated within the project by KU Leuven | .docx, .pdf, .tif, .pptx, .xls | <20 GB |
| VR/AR 4D reconstructions of the activation maps | Generated within the project by KU Leuven and a partner "Digital Arts and Entertainments" | .mp4 | 100 MB x 100 = 10 GB |
| Metadata | Generated within the project by KU Leuven (see later) | .txt/.docx | 100 MB |

## 3. Ethical and legal issues

**3.1. Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to the file in KU Leuven's Record of Processing Activities. Be aware that registering the fact that you process personal data is a legal obligation.**

Yes.

File reference: GDPR questionnaire: File G-2021-4324 (accepted by PRET of KU Leuven 03/02/2022), S66094 accepted by UZ Leuven EC on 25/03/2022.

Within this project, we will collect medical data in healthy volunteers and cardiac arrhythmia patients at UZ Leuven. The type of data include:

- Identification information (names)
- Personal traits (e.g. age, gender)
- Physical traits
- Ultrasound scans performed with low frame rate, commercial ultrasound systems
- Ultrasound scans performed with a research high frame rate ultrasound system
- MRI Diffusion Tensor Images of the heart
- Intracardiac activation maps (obtained during the ablation procedure)

Moreover, we will reuse ex-vivo porcine hearts from other ongoing animal experiments, which regularly take place in the department of cardiovascular sciences at KU Leuven.

**3.2. Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, add the reference to the formal approval by the relevant ethical review committee(s).**

Yes, we will use echocardiographic images which provide information about the subject's cardiac function, and in that sense also provide information about his/her health.

Reference: GDPR questionnaire: File G-2021-4324 (accepted by PRET of KU Leuven 03/02/2022), S66094 accepted by UZ Leuven EC on 25/03/2022.

**3.3. Does your research possibly result in research data with potential for tech transfer and valorisation? Will IP restrictions be claimed for the data you created? If so, for what data and which restrictions will be asserted?**

Yes.

All imaging technologies to be developed in WP1 are of potential interest for vendors of echocardiographic systems. Indeed, improved functional imaging of the heart remains a major target for industry because of the immense clinical need. In particular, an extension to 3D of the shear wave elastography technology, the cardiac 3D deformation imaging, and the fiber orientation imaging technique.

Next to the ‘direct’ imaging technology described above, the inference of tissue parameters through modeling (i.e. ‘indirect imaging’) is highly relevant for the clinic, and as such could be valorized.

For this reason, we aim to restrict access to these until its valorization potential is either realized or abandoned.

**3.4. Do existing 3rd party agreements restrict dissemination or exploitation of the data you (re)use? If so, to what data do they relate and what restrictions regarding reuse and sharing are in place?**

No.

## 4. Documentation and metadata

**4.1. What documentation will be provided to enable understanding and reuse of the data collected/generated in this project?**

In general, each researcher involved in data collection in the project will provide a detailed description of data acquisition in paper of electronic lab notebooks. Protocols will be stored as .docx files.

1. Ultrasound/MRI images and recordings (of phantoms/ex vivo tissues/patients):

For image and video data, text-based readme files will be developed that describe what is captured within these visual datasets, how they can be opened, and why they were collected. Whenever necessary, additional files containing supplementary information (axis, image dimensions, acquisition parameters) will be generated and stored in the same folder. Images will be sorted in folders, whose names will follow a clear structure defined in advance and maintained in a specific .txt file.

All patient data will be coded. Personal data, ECG information and conventional ultrasound images will be stored on a UZ Leuven server and in KWS (Klinisch Werkstation). Metadata are automatically captured when data are transferred in KWS.

2. Electrocardiograms:

Metadata (e.g. timestamps, duration of different ECG intervals, QRS duration) are automatically captured when data are transferred to KWS.

3. Ultrasound simulation data:

Each simulation data will be accompanied by a .txt file with a clear description of what the data represents and how it has been generated. Moreover, the input files related to the simulations will be stored in the same folder. The folder names will follow a clear structure defined in advance and maintained in a specific .txt file.

4. Ultrasound image acquisition and processing software:

The following documentation will be provided:

- structured comments within the source code
- a simple demonstration script that shows how the software can be used
- a text-based readme file that describe the software dependencies and provides installation and compilation instructions
- a tutorial document (PDF file describing the high-level design and organization of the software packages)

This is also valid for other type of software developed within Icarus.

5. Inverse mathematical models: Same principle as 4. above.

6. Manuscripts (and associated figures), protocols, presentation and reports:

These will be stored in a specific folder (publications, protocols, presentations, reports, meetings, etc) in a directory shared within the project members (KU Leuven OneDrive). A file naming convention will be defined in a .txt file.

7. Lab books:

Electronic lab books (and scanned PDFs of manual lab books) will be stored in a specific folder in a directory shared within the project members (KU Leuven OneDrive).

**4.2. Will a metadata standard be used? If so, describe in detail which standard will be used. If not, state in detail which metadata will be created to make the data easy/easier to find and reuse.**

For most of our data (e.g. raw ultrasound data, ECG recordings,...), there is no standard metadata format available. Therefore, relevant metada will be documented by the project members.

Wherever applicable, the Dublin Core Metadata will be used. The original DCMES (Dublin Core Metadata Element Set) Version 1.1 consists of 15 metadata elements: Title, Creator, Subject, Description, Publisher, Contributor, Date, Type, Format, Identifier, Source, Language, Relation, Coverage, and Rights. Each file will thus contain information on study design, sampling methodology, variable-level detail, and all information necessary for a secondary analyst to use the data accurately and effectively.

Moreover, we will closely monitor MIBBI (Minimum Information for Biological and Biomedical Investigations) for metadata standards that are more specific to our data.

## 5. Data storage and backup during the project

**5.1. Where will the data be stored?**

Patient data: A unique code/serial number will be assigned to each participant based on the timing of their inclusion in the study. The subject's name or other identifiers will be stored separately from their research data and replaced from the aforementioned unique code to create a new identity for the subject. When data are coded, there continues to be a link between the data and the individual who provided it. The patients' names and patients' identifiers will be kept in a password protected file, on a secure server with restricted electronic and physical access: The key file will be kept separate from the pseudonymized data, stored on a UZ Leuven server, implying data storage is safe (with back-up & password protected). This key file will only be available to the KU Leuven research team of the project. The personal data will be deleted after 25 years. The signed Informed Concent forms of the participants with a clear description of the study protocol will be stored as hard copies in a locked drawer accessible only to the researchers of the study.

Patient data (once pseudonymized) will be copied to a shared folder on the Kulak servers (password protected) so that it can be used by the collaborators in the group of Prof. Dierckx (who do not have a direct access to the hospital servers).

US imaging experiments and simulations: The experimental ultrasound acquisition data (non-patient data) and simulation data will be stored in the servers of UZ Leuven, in the folder belonging to the laboratory of Cardiovascular Imaging & Dynamics (team of Jan D'hooge). The servers are managed by the ICT department of UZ Leuven.

Simulations for inverse and forward mathematical model results: will be stored in a shared folder on the Kulak servers.

The restricted data (i.e. the developed softwares) will be stored in Gitlab Enterprise repositories at KU Leuven, which are managed by the ICTS services of KU Leuven. The data will be shared between the different members of the project.

Other type of data (electronic lab notebooks, manuscripts, figures, reports, protocols, presentations, etc) will be stored in a shared folder (shared between members of the project) on OneDrive for Business (KU Leuven).

**5.2. How will the data be backed up?**

Gitlab repositories at KU Leuven and OneDrive for Business are managed by the university ICT staffs

and automatically backed-up. So too are the servers managed by UZ Leuven (patient data) and KU Leuven Kulak (J-Drive).

**5.3. 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 total required storage capacity was estimated to be ~3.5 TB.

The Gitlab repositories have a space of 1 GB per user (extendable), which is sufficient for the software that is expected to be generated within the project.

Additionally, the Cardiovascular Imaging and Dynamics Laboratory has 60 TB of storage on the UZ Leuven servers, and this capacity can be extended if required. The cardiac Modelling team at Kulak has access to several TB of storage on shared disks (J drive).

Each KU Leuven member benefits from 2TB on OneDrive for Business (extendable), which should be more than sufficient for the data that will be stored on OneDrive (see table in section 2).

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

The cost for storage of data on the hospital servers is 1200 euros/TB/year (including back-up).

The cost for storage of data on the Kulak servers is 519 euros/TB/year (including back-up).

Storage on OneDrive and Gitlab is expected to be free of cost due to the sufficient available space provided by KU Leuven.

The costs have already been factored into the consumables budget for this project (5000 euros have been budgeted for data management). The storage capacity will gradually increase as the project progresses.

**5.5. Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?**

Data security is provided by the ICT departments of KU Leuven and UZ Leuven. They will maintain the servers on which the data is stored, backup the data regularly, and fight against any malware attacks. The UZ Leuven servers offer secured access to folders.

In the case of the restricted data, access will be password protected and be provided to the project members by default. Any additional users who wish access to the restricted data will receive it only if approved by all project investigators (Prof. Dierckx, Prof. D'hooge and Prof. Ector). This is also true for the working copies of the shared data that will be saved on the university servers.

## 6. Data preservation after the end of the project

**6.1. Which data will be retained for the expected 10 year period after the end of the project? If only a selection of the data can/will be preserved, clearly state why this is the case (legal or contractual restrictions, physical preservation issues, ...).**

All data listed above will be retained for the expected 10 year period after the end of the project. We only make an exception for simulation results (forward and inverse models, or ultrasound image simulations), because the simulation output can be very heavy (several GB per file), and can be easily recreated from the source codes (which will be conserved).

The patient data will be stored for 25 years as required by the hospital policy.

**6.2. Where will these data be archived (= stored for the long term)?**

The data will be stored on the university's central servers (with automatic back-up procedures) for at least 10 years, conform the KU Leuven RDM policy. Patient data will be stored in the UZ Leuven PACS for 25 years conform to the hospital policy.

**6.3. What are the expected costs for data preservation during these 10 years? How will the costs be covered?**

Storage on KU Leuven Archiving disks (K drive): the software data, processed data, and all the data stored on Git Lab, OneDrive, and on the J: shared drives will be transferred to an archive (K drive) for long-term storage. In addition, a copy will be made on an external hard drive, which will be stored in the office of Prof. D'hooge. This ensures that the data is easily accessible both for Kulak project members (K-drive) and for UZ Leuven project members (hard drive copy). The storage cost is 99.55 euros/TB/year.

Storage of patient data: patient data will remain on the hospital servers (offline storage) for 25 years. The estimated storage cost is xx.

These costs have been factored into the consumables budget for this project (5000 euros have been budgeted for data management).

## 7. Data sharing and re-use

### 7.1. Are there any factors restricting or preventing the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions or because of IP potential)?

The software generated in this project may have valorization potential and therefore, we may restrict access to this software until this valorization potential has been exploited.

A decision on valorizing these data will be made by the end of the project, and if the valorization potential is not there, these data will be made available.

The patient data will not be shared on public repositories, in agreement with the GDPR policy.

### 7.2. Which data will be made available after the end of the project?

Data uploaded on journals as supplementary materials, and softwares (forward and inverse mathematical models, ultrasound image acquisition & processing softwares) on GitLab that we choose to make public (if not valorized).

### 7.3. Where/how will the data be made available for reuse?

- In a restricted access repository
- Upon request by mail
- Other (specify):

Annotated data are uploaded to KULeuven repository RDR and/or as supplemental material to publications, if requested by the journal.

### 7.4. When will the data be made available?

- Immediately after the end of the project
- Upon publication of the research results

Data shared via supplemental material is shared upon publication. Other data are shared after the valorization potential has been fully explored.

### 7.5. Who will be able to access the data and under what conditions?

Synthetic data (e.g. simulation output) can be accessed freely, but requires citation of the source and paper if used by other parties.

The data uploaded on the KU Leuven repository RDR will be available to third parties under restrictions as defined by members of the project (e.g. email request).

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

The costs for data sharing with journal platforms is covered by the publication fee. For data stored at RDR, the costs are covered by KU Leuven.

## 8. Responsibilities

### 8.1. Who will be responsible for the data documentation & metadata?

For patient data: Konstantina Papangelopoulou (PhD student of Prof. Ector)

For simulations and source codes, experimental acquisitions: Nathan Dermul (PhD student of Prof. Dierckx) & Sophie Heymans (postdoctoral researcher of Prof. D’hooge)

Each researcher will be responsible for documenting his/her own laboratory notebook, publications, research output, presentation materials, etc.

### 8.2. Who will be responsible for data storage & back up during the project?

For patient data: Konstantina Papangelopoulou (PhD student of Prof. Ector)

For simulations and source code + experimental acquisitions: Nathan Dermul (PhD student of Prof. Dierckx) & Sophie Heymans (postdoctoral researcher of Prof. D’hooge)

Data backup is automatically performed by ICT of KUL and UZ Leuven.

### 8.3. Who will be responsible for ensuring data preservation and sharing?

Prof. Dierckx will be responsible for data preservation and sharing.

### 8.4. Who bears the end responsibility for updating & implementing this DMP?

Sophie Heymans bears the overall responsibility for updating the DMP, while Prof. Dierckx bears the responsibility to ensure that the DMP is correctly implemented.