E.D.G.A.R. User’s Guide

Version 1.0

Last update: July 27, 2015

E.D.G.A.R. - Experimental Data and Geometric Analysis Repository

Authors:
Wilson W Good (wgood@sci.utah.edu),
Kedar Aras (kedar.aras@gmail.com),
Jess Tate (jess@sci.utah.edu),
Brett Burton (bburton@sci.utah.edu, and
Rob S. MacLeod (macleod@sci.utah.edu)

Scientific and Computing Institute (SCI)
www.sci.utah.edu

Nora Eccles Harrison
Cardiovascular Research and Training Institute (CVRTI)
www.cvrti.utah.edu

The University of Utah

Support for this project came from:

The Nora Eccles Treadwell Foundation

To reach the EDGAR team please contact:

edgar@sci.utah.edu

Contents

1 What is E.D.G.A.R.?

The Experimental Data and Geometric Analysis Repository (EDGAR) is a free, web-based repository hosted by the SCI Institute at the University of Utah. The purpose of EDGAR is to share and collate electrocardiological data, specifically for the validation and advancement of electrocardiography imaging (ECGI) problems amongst a worldwide consortium of academic institutions. This repository is coordinated through the Consortium for ECG Imaging (www.ecg-imaging.org) and we encourage all users of the resource to acknowledge the role of this group in their research.

2 Introduction

2.1 Overview

The EDGAR project provides free and open access to experimental and clinical ECGI data acquired and prepared by academic and commercial centers around the world. As a vehicle for the Consortium for ECG Imaging (CEI), the aim of the EDGAR project is to catalyze scientific progress and understanding in the broad field of electrocardiology and, more specifically, forward and inverse problems—ECG Imaging. Participants in the project will be able to freely implement their particular forward/inverse solutions, offer their own strengths to other participants, and evolve existing techniques or invent new, promising techniques. By providing a wide range of time signals, geometry, images, transfer matrices, and documentation, we allow users to identify and access data that are related to their own research ventures. By providing a wide range of data, we support a high degree of specialization within the field, as, for example simulation experts need not develop their own experimental or clinical techniques for data acquisition and the data specialists can benefit from th expertise of image based modeling specialists. The aggregation of experimental and clinical data from various centers will allow modelers to glean from a greater and more diverse data pool adding to the robustness and accuracy of the intended model.

The ECG has been the primary mode of diagnosing cardiac arrhythmias and other cardiac dysfunctions for the better part of a century despite its limited spatial resolution and substantial ambiguity of data interpretation. The diagnostic limitations in localizing sources of irregular electrical behavior from the ECG is also a result of the attenuating, smoothing, and distorting effects of the body through which cardiac currents flow. Body Surface Potential Mapping (BSPM) provides increased resolution through additional spatial coverage while geometric models of the torso provide a means to compensate for this smoothing. The combination of these additional information sources to improve recovery of cardiac electrical activity is the goal of recent ventures into the electrocardiographic inverse problem.

Being able to accurately identify electrical activity in the heart from the ECG is extremely useful in cardiac diagnostics because it is non-invasive, painless, inexpensive, and capable of continuous sampling. Unfortunately, the inverse problem is ill-posed and may be non-unique, leading to poorly conditioned systems of equations to solve. As one result, small variations in the model or the ECG signals can result in unconstrained fluctuations in the cardiac source description.

2.2 Intended Audience

The target audience for EDGAR is primarily researchers with at least some facility in engineering, particularly those involved in computational cardiology or simulation research. We assume uses of the archive will either have their own software for reading and manipulating the files or that they will make use of some of the tools described on the CEI website (www.ecg-imaging.org). Members of the EDGAR team are also available to provide assistance for specific questions. For those interesting in learning more about the field and the tools required to carry out ECGI, please also see the workshop and learning resources listed on the CEI website.

2.3 Goals and Objectives

The goal of EDGAR is to establish a communal forum for the sharing and distribution of cardiac electrophysiology data and geometric models intended to solve forward-inverse problems. The motivation in creating and hosting a database like EDGAR is to make data accessible to multiple experts with complementary skills in a collaborative setting. The collective goal is to continue to improve the accuracy, utility, and impact of ECGI. It is our hope that EDGAR will be used to create increasingly accurate and robust cardiac models that will produce novel insights into the physiology of the heart while improving the manner in which we recognize and treat disease states in a clinical setting.

The sources of data for EDGAR include physiological experiments, both with animals and human subjects, as well as simulated signals generated from electrophysiological models of the heart. Such data are challenging to acquire and the contribution of data donors should be clearly recognized and acknowledged by EDGAR users. We strive to make these data sets as complete and accurate as possible, with extensive documentation about the experiments, the means used for acquisition, and the post processing applied. Some data sets are more complete than others and all experiments are compromises; complete documentation should ensure the best use possible of these valuable resources.

3 Organization and Expected Interactions

Figure 1 shows the organization of Edgar data at the highest level, consisting of a set of *metadata* and documentation for each experiment along with associated data files. All the data contained within the database will be organized into the following five categories, which may expand in the future as needs arise:

1. Images
2. Geometric model(s)
3. Registration information
4. Time signals
5. Transforms (or transformation matrices)

A second, orthogonal organization of the EDGAR data is through a hierarchy that starts with an **experiment**, assumed to take place on a single day with a name that includes the experiment date (or a pseudo-date, when necessary, to protect patient confidentiality) and other identifiers specific to the organization that produced it (*e.g.,* the prefix “RSM” indicates an experiment from the MacLeod team at the University of Utah). Each experiment encompasses a sequence of **interventions**, which are based on any structure that is meaningful for the specific case. For example, in animal experiments, such interventions could capture external manipulations such as pacing sites and rates, ischemia, or the application of a drug to alter physiology of the preparation. In clinical studies, interventions could be organized around, for example, a sequence of paced mapping or variable rates from the same site. Within each **intervention** is a series of **runs** that are continuous recordings of signals of variable duration—the key requirement being continuous sampling of signals at a stable sampling rate. Examples of run data may be continuous sampling of several seconds or a single beats extracted from longer recordings.

The two organizational structures described above are intertwined in that different elements of the experiment/intervention/run structure each have the possibility to contain different sets of data. Table 1 summarizes these different associations, described in detail as follows:

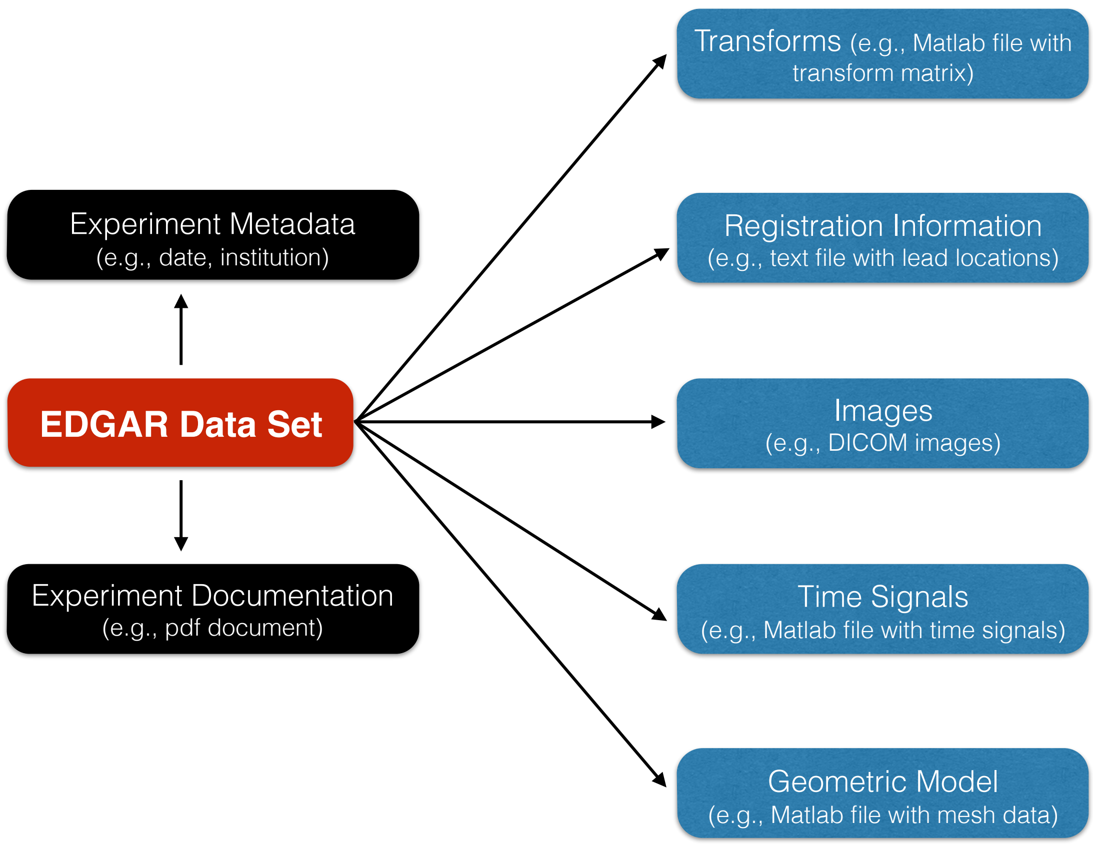

Figure 1: Layout of the Edgar system by types of data. This view consists of metadata and associated data files, each of which contains a specific type of data, and have their own formats. The associated data files can be downloaded individually, on an intervention basis, or as an entire experiment.

Experiment: may have one or more sets of images, must have at least one geometric model/mesh that captures the anatomy, and the locations of the sample sites for the time signals. Each experiment will also typically have one set of registration information and may have one or more transforms.

Intervention: may have its own images, geometric model, and registration information but it must have at least one **run**. The runs that make up an intervention belong in some organizational sense together and typically represent a sequence of runs with a varying level of some control parameter. Typical examples are a set of paced beats, each initiated from a different locations or a progression of an intervention such as ischemia or the level of a drug.

Run each **run** consists primarily of an associated set of time signals, one for each of the measurement sites. In additional, a run may have its own images, geometric model, and registration information, although this will not be the typical case. In cases in which the run contains its own images or additional files, they will stored on an interventional basis with an indication to what run they are associated with. Each run may also contain signals that have been interpolated or otherwise processed rather then directly measured. Thus multiple time signals may be associated with a single **run**.

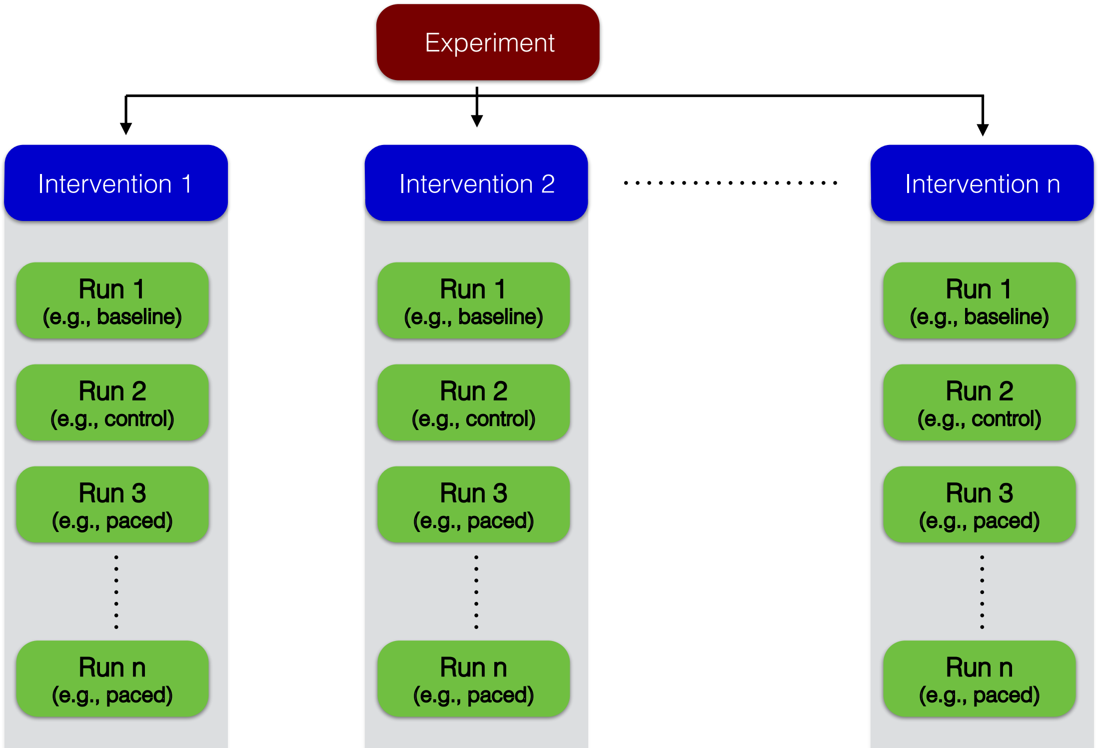

Figure 2: Layout of the Edgar system by sequential components. This view consists of experiments, interventions, and runs, each of which can contain elements from the data organization in the previous figure.

4 File Formats

4.1 Metadata

The metadata aspect of the EDGAR system is the most flexible and hence the most challenging to prescribe. The details of this section will evolve rapidly as we gain experience. Essentially, the metadata consists of searchable terms is used to identify and find datasets that conform to a key set of parameters (*i.e.,*Species, Study Type, or Data Origin). The current list of such terms is as follows:

Data origin: the physical location of the source of the data, typically the site of the experiment or clinical study captured by the data.

Experiment Date: The date at which the experiment was undertaken. Can be a pseudo date in order to protect patient confidentiality.

Species: the species of the study subject, typically a large mammal or human patient.

Study type: The overall goal or nature of interventions included in the experiment. Typical examples are ischemia or ectopic pacing.

Level of Hierarchy	Potential Contained Data	Example/Justification
Experiment	Medical Images, Registration (Reg), Documentation, Meshes	**Images** – Sequences acquired are used for all interventions **Registration** - transformed geometry is used for all interventions **Documentation** – Experimental protocol and ReadMe docs that give upper level descriptions of the dataset **Meshes** – Non-conforming meshes
Intervention	Time Series (TS), Meshes, Images, Registration (Reg), Transforms(Fwd/Inv)	**Time Series** – Data acquired per intervention. **Meshes** – geometry is adapted in a run specific manner **Images** – Sequences acquired are used on an intervention basis **Registration** – transformed geometry applies only at the intervention level **Transforms** – Projections of run specific data onto mesh data

Table 1: Example cases of where certain data can be found/placed based on the layout of the experiment and the nature of the run data.

4.2 Medical Imaging Files

Some EDGAR data sets will include medical images from the torso and/or heart, which will form the basis for geometric models. The primary format for these files is the Digital Imaging and Communications in Medicine (DICOM) (http://dicom.nema.org), which is a very well documented and supported international standard. Software to support this format will come largely from commercial and open source tools. A secondary format for image data will be the Nearly Raw Raster Data (NRRD)
http://teem.sourceforge.net/nrrd/, which is an open source format with support from the TEEM library http://teem.sourceforge.net and other tools based on this library, *e.g.,* Seg3D
http://www.sci.utah.edu/cibc-software/seg3d.html, 3DSlicer http://www.slicer.org,
or ITKSnap http://www.itksnap.org.

4.3 Segmentation

Segmentation in the context of this project is the process of identifying and differentiating tissue regions in the torso, either by creating surfaces that separate or contain regions or by labeling voxels within the image data. EDGAR provides support for segmentation files in the form of "label masks", which are rasterized, volumetric labels that must be assigned to the original image data. Hence, label masks have the same size and spacing as the image files from which they are derived.

This project supports segmentation files in the NRRD format. The process of performing segmentation can be very challenging and automating segmentation is a major research topic. Tools for performing segmentation include the same selection as above: Seg3D http://www.sci.utah.edu/cibc-software/seg3d.html, 3DSlicer http://www.slicer.org, or ITKSnap http://www.itksnap.org.

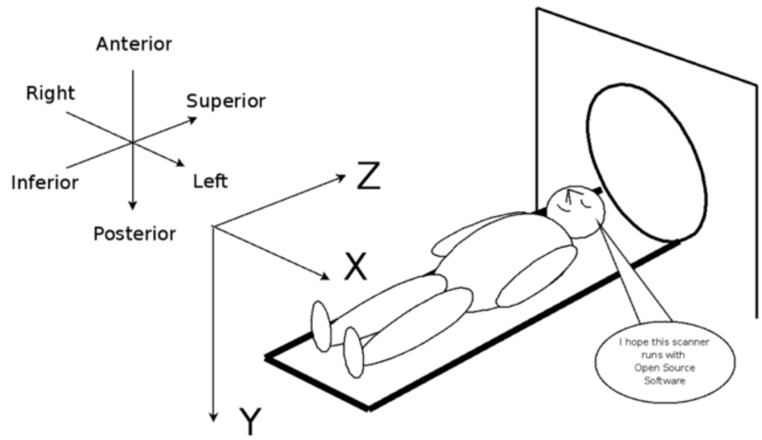

Figure 3: **EDGAR Coordinate system** All EDGAR geometry models assume a body coordinate system, as shown in the figure.

Orientation is a important factor to consider when generating and analyzing geometric models. Because of the great deal of diversity in the origin of the data, some degree of consistency is necessary to facilitate comparison across models. We have adopted and imposed a body reference coordinate system, shown in Figure 3. In order to maintain consistency across the data acquired from various academic centers, we transform all geometry to this axis ordering. We also assume units of millimeters for all geometry models. The origin will not be fixed to any anatomical construct in the geometries as this would be difficult to enforce across a broad and diverse pool of datasets. We assume that the origin of the coordinate system is within the heart or body so as to facilitate visualization.

4.4 Geometric Models

The geometry of interest for forward and inverse ECGI are usually captured by “meshes”, a series of node locations in 3D space connected by a network of edges to form polygons, typically rectangles and triangles in two dimensions and hexahedra and tetrahedra in three dimensions. These meshes, most importantly, form the support for numerical solutions of the equations of interest, *e.g.,* finite element, boundary element methods, and the finite differences method. Meshes can also serve as the basis for visualization of the associated measured or simulated potentials and currents.

Within EDGAR, meshes are most often associated with a single experiment, *i.e.,* a single torso and heart geometry that is maintained throughout the experiment. However, geometry may change over the course of an experiment so that it is possible to link meshes at the intervention and run levels of the file system.

This project supports mesh files in MATLAB file format and, in some cases as ASCII/text files using specific organization described below. We anticipate modifications to these formats but will always strive for simplicity and portability.

4.4.1 MATLAB geometry file support

EDGAR supports .mat files from MATLAB as long as they are organized according to the following guidelines:

1. Each separate surface is in the form of a structure (see the MATLAB documentation on structures). To include multiple surfaces requires an array of structures.

2. Within each surface structure, the following fields contain the geometry:

   - *.pts or .node* contains the node locations, usually in a $3 \times N$ array (although *map3d* will check and accept either $3 \times N$ or its transpose, $N \times 3$), where N is the number of nodes.
   - *.fac or .face* contains the triangle connectivities, usually in a $3 \times M$ array, where M is the number of triangles. **Notes: The numbering of nodes for connectivities begins as zero and not one**. Order of nodes is not formally constrained but we encourage adopting a convention in which the order is counterclockwise when viewed from the outside of the surface defined by the triangles.
   - *.seg or .edge* contains the line segment connectivities,
   - *.tet, .tetra, or .cell* contains tetrahedral connectivities. Each tetrahedron may contain an optional fourth value which is the group number for the tetrahedron, typically associated with a tissue type and eventually an electrical conductivity value in simulations. Order of tetrahedra in the file is not meaningful except for internal bookkeeping.

To prepare a structured .mat file is very simple, for example using the following commands:

```
>> geom.pts = ptsarray;
>> geom.fac = facarray;
>> save mygeom.mat geom
```

where `ptsarray` is a $3 \times N$ array defined to contain the node locations, `facarray` is a $3 \times M$ array of triangle connectivities, and `mygeom.mat` is the name of the resulting .mat file.

4.4.2 Text file formats

Points (.pts) file The characteristics of a .pts file are as follows:

1. ASCII file, no special characters permitted;

2. Each line contains one triplet, ordered as x, y, and z values; one or more spaces between values, which are assumed to be real, floating point numbers;

3. the order of points in the file is the implicit order of the nodes in the geometry; connectivities are based on this ordering. This order must also match, one-to-one, the order of the time series in data runs.

Triangle (.fac) files The characteristics of a .fac file are as follows:

1. ASCII file, no special characters permitted;

2. Each line contains a triplet of integer values pointing to the nodes of the geometry. **Node numbers begin at 0 not 1!**;

3. The order of triangle vertices (nodes) is not strictly controlled, however, it is recommended that order reflect a common convention in graphics—a counterclockwise sequence of vertices when viewed from the **outside** of the triangle;

4. Each line may also optionally contain a group number as a fourth element, which is used to indicate tissue type from the original image data.

5. Order of triangles in the file is not meaningful except for internal bookkeeping.

Tetrahedra (.tet) files The characteristics of a .tet file are as follows:

1. ASCII file, no special characters permitted;
2. Each line contains a quadruple, (list of four) of integer values pointing to the nodes of the geometry. **Node numbers begin at 0 not 1!**;
3. The order of tetrahedra vertices (nodes) is not strictly controlled, however, it is recommended that order reflect a common convention in graphics—a counterclockwise sequence of face vertices when viewed from the **outside** of the element;
4. Each line may contain an optional fourth value which is the group number for the tetrahedron, typically associated with a tissue type and eventually an electrical conductivity value in simulations;
5. Order of tetrahedra in the file is not meaningful except for internal bookkeeping.

4.5 Registration Files

Registration is the process of bringing a series of geometric entities into the same coordinate system such that signals acquired during the experiment can be spatially arranged in the desired coordinate space. For example, in an animal experiment, this information may come using a mechanical digitizer; in a clinical electrophysiology study, the information can be contained in a file generated from an electroanatomical mapping system and saved during the procedure. Each source will have its own file formats and assumptions so that standardization is particularly challenging. To provide the necessary information, EDGAR will require a text description in the documentation section that explains the necessary details.

As an example from a typical animal experiment at the University of Utah, registration between the heart and torso geometry happens as follows:

1. Registration is by means of a mechanical, three-dimensional digitizer that captures coordinates as landmarks in the torso tank and the heart. Each landmark/registration point is the mean of three replicates and is stored in a simple text file for subsequent post processing. We typically collect approximately 10 locations on the torso tank and 6–15 locations on the heart surface, usually points from an electrode sock or from the entry locations of intramural recording needles.
2. Locations of the heart landmark points are captured from subsequent imaging of the heart or from known locations on a geometric model of the epicardial sock and/or the needle electrodes.
3. We register the heart and torso tank using a rigid body (procrustes) method that optimizes translation and rotation based on minimizing the sum of the squared differences between landmarks in the torso tank and the heart coordinate reference frame. The result is a complete geometric description of the heart and torso in the coordinate system of the torso.
4. Included with the experiment is a listing of the measured point coordinates for all the landmarks and electrode locations as well as a fully registered geometric model of the torso tank, the heart, the epicardial sock electrodes, and the intramural needle electrodes. Also included are the registration matrices applied in standard homogeneous coordinates (https://en.wikipedia.org/wiki/Homogeneous_coordinates).

Such descriptions of the geometric source data and subsequent processing should be included in a text description, *e.g.,* ReadMe.tex file or pdf file attached to the data. Registration techniques vary significantly between institutions so it must be understood that the method employed by the institution should be concisely detailed in the documentation.

4.6 Forward/Inverse Transforms

This section of the EDGAR data set may contain the transformation matrices that project signals from a cardiac source descriptions to the body surface, *i.e.,* one form of the forward solution, as well as the corresponding inverse transform matrix, a solution to the inverse problem. The formulations used to create these transformations are rather diverse and we will include descriptions of these methods and assumptions as part of the documentation. The transformation matrices should be constructed such that the matrix provides 1 to 1 mapping between the relevant time series data and the projected surface. The row and column composition should be made compliant with the target data in a plug and play type manner so that a user may assume that for any forward transform matrix

$$\vec{\Phi_B} = Z_{BH}\,\vec{\Phi_H}, \tag{1}$$

where Φ_B is a vector of body-surface potentials, Φ_H is a vector of heart potentials or activation times, and Z_{BH} is the transformation matrix. Inverse transform matrices are similarly formulated as

$$\vec{\Phi_H} = Z_{HB}\,\vec{\Phi_B}. \tag{2}$$

Forward and inverse transformations can be formulated and analyzed using the tools provide in the SCIRun computational platform with the Forward/Inverse toolbox distributed with the software. SCIRun can be downloaded from http://www.sci.utah.edu/cibc-software/scirun.html.

4.7 Time Series Data

The standard format for time signals is MATLAB, which, while not an open source language, is supported by open source file readers and even a reasonably functional equivalent (http://www.gnu.org/software/octave/). As with the geometry files, the format is a structure with a field ".potvals", that contains an $N \times M$ array, where N is the number of data channels and M is the number of time instants. There are additional fields permitted in the structure, the complete list is as follows:

- *.potvals, .data, .field, or .scalarfield* scalar values as $N \times M$ array, where N is the number of data channels and M is the number of time instants.
- *.unit* the type of units for the data, "uv" for microvolts, "mv" for millivolts and "V" for volts.
- *.label* the name of the time series. This is optional, but is useful in identifying the time series, particularly from a multi-time-series file.
- *.fids* a structure (or array of structures) containing fiducial time markers for the time series.

Note that only the 'potvals' field is required. A MATLAB array may be one instance of these fields, a cell, or struct array of them, or simply an $N \times M$ array of data.

The commands to make a suitable .mat file are very easy in MATLAB, for example, to load 128 channels of time signals in units of millivolt from an array `sockinfo` to a file called `mysockdata.mat`

```
>> sockdata.potvals   = sockinfo(1:128,:);
>> sockdata.unit = 'mV'
>> sockdata.label = 'A set of cool sock data'
>> save mysockdata.mat sockdata
```

5 REDCap Interface

The EDGAR metadata storage scheme is implemented in REDcap, a secure database, web application specialized for scientific research, and hosted at the University of Utah (www.redcap.org). Typical interaction with the EDGAR front end includes searching for data sets based on specific metadata criteria (*e.g.,* species, study, date, etc.) or alternately browsing through all the available data sets. These search terms will filter the results based on the labeling information stored within the metadata, the results will be returned on a experiment basis, and subsequent browsing will be needed to navigate to intervention level data. Each of these metadata entries has a pointer to a zip file containing the files of interest. The user then sees an indication of the size and number of contained files and chooses the desired dataset(s) and component(s). To initiate the search, each user must establish an identity on the RedCap server based on an email address but without any need for a password. Subsequent access follows with the email address alone.

5.1 Web Interface

EDGAR has a web interface (www.ecg-imaging.org/edgar) available through the Consortium for ECG Imaging (CEI) website (www.ecg-imaging.org/) that enables easy, convenient, and efficient access to data. Demographics and logistics about the user are gathered upon first applying to access data on EDGAR, and all such data will remain secure and available only to document usage of the repository and any associated grants or funded research for which the data are used.

Each set of data also include a desired form of acknowledgment for any subsequent use of the data. Such acknowledgment will typically be a sentence to be included in the appropriate sections of publications or proposals.

5.2 New Users

There are user roles within EDGAR: uploading data and downloading data. Each user will be required to agree to a legal statement ensuring that all data files are handled in a compliant manner.

Understandably, EDGAR simply cannot exist without the influx of data and thus a reliable and consistent filing scheme must be implemented. The screening, preparing, and filing of data to the RedCap site will be the responsibility of the SCI Institute. Research groups wishing to upload data to EDGAR will receive an upload request form that includes details of the intended hierarchy and layout of the data as well as an agreement to the EDGAR terms of use.

Each user wishing to download data must agree to the EDGAR access terms when they create an account. The main goals of this agreement are to ensure compliance with all relevant local regulations for the use of human or animal data and to encourage acknowledgment of the data sources in subsequent publications.

The fields and types of information collected in creating a user account are:

- *Name*
- *Email Address*
- *Institution*
- *Location*
- *Purpose for Downloading data*
- *Interest in signing up for mailing list*
- *Supporting Grant(s)*

5.3 Searching

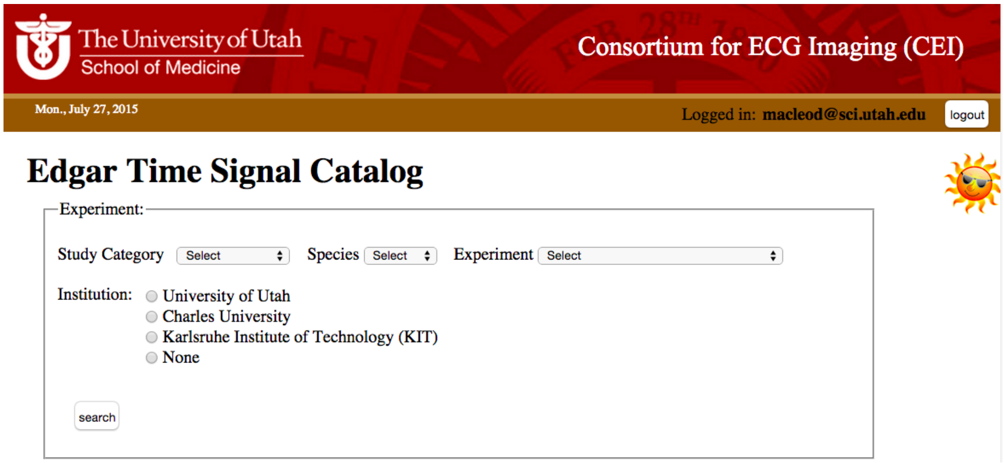

Figure 4: **Example of the user interface to EDGAR through the REDCap datbase.** The image shows the top level of the search function with the main metadata fields exposed. Details of the interface will change with time; the image here may fall out of date at any time.

REDCapSearch-list The search parameters for EDGAR are quite simple, limited to those parameters seen in Figure 4. For each search category, a series of check boxes allow the user to select the specific data to use without having to download any experiment. An alternative check box will allow the user to download all of the data contained within a single experiment. Further detail concerning the experimental procedure and contents can be seen by directly downloading the 'Documentation' folder of each experiment.

Experiment (click to view experiment files)	Description
PSTOV-12-07-29	This directory contains the data for the pacing localization experiments done in Prague by Petr Stovicek. It consists on a series of body surface recordings from a heart being paced with a catheter device. The pacings were done on the endocardial surface of the (healthy) ventricles of the subjects. Pacing sites are stored on an interventional basis. These datasets contain surface geometries of the heart and torso. The latter geometry consists on the Dalhousie torso morphed to fit the patient's on axial CT scans limited to the chest.
PSTOV-12-07-27	This directory contains the data for the pacing localization experiments done in Prague by Petr Stovicek. It consists on a series of body surface recordings from a heart being paced with a catheter device. The pacings were done on the endocardial surface of the (healthy) ventricles of the subjects. Pacing sites are stored on an interventional basis. These datasets contain surface geometries of the heart and torso. The latter geometry consists on the Dalhousie torso morphed to fit the patient's on axial CT scans limited to the chest.
PSTOV-12-07-28	This directory contains the data for the pacing localization experiments done in Prague by Petr Stovicek. It consists on a series of body surface recordings from a heart being paced with a catheter device. The pacings were done on the endocardial surface of the (healthy) ventricles of the subjects. Pacing sites are stored on an interventional basis. These datasets contain surface geometries of the heart and torso. The latter geometry consists on the Dalhousie torso morphed to fit the patient's on axial CT scans limited to the chest.

Figure 5: **Sample listing of interventions in one sample experiment.** Selecting a source of data then reveals a list like this of all the separate experiments from that source.

Search refinement beyond the experiment level can be attained through manually browsing entries on an experiment by experiment basis. Refined searching follows clicking on the data entry under the 'Experiment' column. In Figure 4, for example, deeper searching would be done by selecting the “Charles University” Institution label. This would then provide a brief description of the experiment(s) available from that source, as shown in Figure 5.

- Docs.zip
- FwdInvTransforms.zip
- Interventions
  - interventionLeftVentPace1.zip
  - interventionLeftVentPace11.zip
  - interventionLeftVentPace12.zip
  - interventionLeftVentPace13.zip
  - interventionLeftVentPace14.zip
  - interventionLeftVentPace15.zip
  - interventionLeftVentPace16.zip
  - interventionLeftVentPace17.zip
  - interventionLeftVentPace18.zip
  - interventionLeftVentPace2.zip
  - interventionLeftVentPace3.zip
  - interventionLeftVentPace4.zip
  - interventionLeftVentPace5.zip
  - interventionLeftVentPace6.zip
  - interventionLeftVentPace7.zip
  - interventionLeftVentPace8.zip
  - interventionLeftVentPace9.zip
  - interventionRightVentPace1.zip
  - interventionRightVentPace2.zip
  - interventionRightVentPace3.zip
  - interventionRightVentPace4.zip
  - interventionRightVentPace6.zip
- Meshes.zip

Figure 6: **Contents of a single intervention.** Selecting an intervention then reveals a list of runs and associated data files within that intervention.

To drill further into the details of an experiment, selecting an intervention from the list (*e.g.,* Figure 5) then reveals a list of runs and the associated geometry, images, transforms, *etc.,* files for that intervention, as shown in Figure 6. Note that all data that is not local to the intervention will only be available at the experiment level, *e.g.,* geometry, images, and documentation that are common to all interventions. Each link in this page is to a zipped archive containing the indicated data.

5.4 Downloading data

Data stored on EDGAR is accessible in a hierarchical manner, meaning all files that exist downstream from the selection point will be included in the download. Therefore, in order to get data associated with an intervention, one may have to search for the experiment that contains the intervention and download some files from the experiment, then proceed to the intervention level to download the details, especially the time signals. The downloadable data at the intervention level will only include the data associated with that intervention.

Downloadable data on the **’Experiment’** level:

- **Time Series Data** - All time series data for the experiment.
- **Mesh Files** - All generic and run specific meshes generated for this experiment.
- **Registration** - The registration data for the experiment.
- **Images** - Medical Images gathered for generating the geometry.
- **Forward/Inverse Trasforms** - The solved transformation matrices (Fwd/Inv).
- **Documentation** - Experimental protocol and a description of the data.

Downloadable data on the **'Intervention'** level:

- **Time Series Data** - Times series for only those runs contained within this intervention.
- **Meshes** - The meshes generated for the runs contained within this intervention.
- **Images** - The images that were acquired for this intervention. (Clinical)
- **Fwd/Inv Transforms** - The transformations that were generated for this intervention.

5.5 User Download Agreement

As a user of the data contained within this repository one must agree to the following statement:

"I acknowledge my responsibility to comply with local laws and restrictions pertaining to the use of these data. I pledge to cite the source of the data as stated in the description of the dataset as well as the documentation folder of each dataset."

6 Data Donor Documentation

6.1 Uploading

For the sake of consistency, the data will be uploaded to EDGAR exclusively by the SCI institute to ensure the data is presented in a sensible and consistent manner. Certain measures will be taken such that all of the contained data conforms to a general hierarchy that is broadly applicable to the contained data. To ease processing, standard data formatting will be applied to the data so that it agrees with the descriptions presented in the 'Organization and Expected Interactions' section. Due to the investment necessary to make this possible, we ask that any and all data provided be presented in a concise and orderly manner that is consistent with the definitions provided earlier in this document.

A request form must be filled out before data is sent to the SCI institute. This form will ensure that you understand the usage agreement of EDGAR and will give us the information that we need to appropriately format the files. To request one of these forms send an email to edgar@sci.utah.edu and a form will be sent to you at our earliest possible convenience.

6.2 Provenance

In compliance with the goal of the EDGAR project to share datasets between institutions, contents of a shared data repository that is collected by one academic center should be used by another entity to generate results and findings. It is also fair to make reasonable attempts to identify and give credit to the original source of the data, known as "provenance". There may eventually be an automatic means to track provenance but in the meantime, we will encourage an informal convention that attaches and acknowledges the source of data.

6.3 Donor Agreement

As a data contributor, submission of data to this repository mandates acceptance of the following statement:

"I control access to the data and promise that it was collected in accordance with all applicable rules in my home country (IRB or IUCAC). I have the right and agree to give the data in perpetuity to the CEI for distribution."

7 Questions and Bug Reporting

We are interested to hear how you feel about using EDGAR and to know if there are any features you would like to see/remove. As we all know software is rarely infallible and it certainly helps to get ready feedback from a wide range of end users as most use cases cannot be predicted. It is this that makes your input so valuable and we would like to keep EDGAR as useful and relevant as we can. Please let us know if you have any questions regarding the operation and management of EDGAR and/or the data stored within the repository. If so please forward these comments to: **edgar@sci.utah.edu.**