The Cardiac Atlas Project: Rationale, Design and Procedures

Carissa G. Fonseca[1], Michael Backhaus[2], Jae Do Chung[2], Wenchao Tao[1], Pau Medrano-Gracia[2], Brett R Cowan[2], Peter J Hunter[2], J Paul Finn[1], Kalyanam Shivkumar[3], Joao AC Lima[4], David A Bluemke[4], Alan H Kadish[5], Daniel C Lee[5], and Alistair A Young[2]

[1] Diagnostic CardioVascular Imaging, UCLA,
Suite 3371, 10945 Le Conte Avenue, Los Angeles, CA 90095-7206, USA
{cfonseca,wtao,pfinn}@mednet.ucla.edu
[2] Auckland Bioengineering Institute, University of Auckland,
Private Bag 92019, Auckland 1142, New Zealand
{m.backhaus,jaedo.chung,p.medrano,b.cowan}@auckland.ac.nz
{p.hunter,a.young}@auckland.ac.nz
[3] UCLA Cardiac Arrhythmia Center,
BH-307 CHS, 10833 Le Conte Avenue, Los Angeles, CA 90095, USA
kshivkumar@mednet.ucla.edu
[4] Department of Radiology, Johns Hopkins Hospital,
600 N Wolfe Street, Baltimore, MD 21287, USA
{jlima,dbluemke}@jhmi.edu
[5] Bluhm Cardiovascular Institute, Northwestern Memorial Institute
251 E. Huron Street, Feinberg Pavilion 8-536, Chicago, IL 60611, USA
{a-kadish,dlee}@northwestern.edu

Abstract. The Cardiac Atlas Project (CAP) is a NIH sponsored international collaboration to establish a web-accessible structural and functional atlas of the normal and pathological heart as a resource for the clinical, research and educational communities. An initial goal of the atlas is to facilitate statistical analysis of regional heart shape and wall motion characteristics, and characterization of remodeling, between and within population groups. The two main early contributing studies are the Multi Ethnic Study of Atherosclerosis (MESA) and the Defibrillators to Reduce Risk by Magnetic Resonance Imaging Evaluation (DETERMINE) clinical trial. De-identified image and text data from 2864 asymptomatic volunteers from MESA, and 470 myocardial infarction cases from DETERMINE, are currently available in the CAP database. DICOM images were de-identified using HIPAA compliant software based on tools provided by the Center for Computational Biology at UCLA. Only those cases with informed consent and IRB approval compatible with the CAP were included. Researchers requesting permission to access CAP data can apply through the CAP website (www.cardiacatlas.org). All proposals for data access must be approved by the data contributors, and applicants must sign a data transfer agreement with each study from which data is requested. Software to visualize cardiac images and create 3D mathematical models, developed in the CAP, is available open-source from the website.

Keywords: Computational Atlas, Database, Cardiac, Mapping.

O. Camara et al. (Eds.): STACOM-CESC 2010, LNCS 6364, pp. 36 – 45, 2010.

1 Introduction

Cardiac performance in health and disease is defined across multiple levels of structure and function from molecular and cellular organization to gross anatomy, and can be studied using a diversity of both imaging techniques (MRI, CT, echocardiography, coronary angiography) and non-imaging tools (ECG, blood pressure and cholesterol measurements). Mathematical and computer models can be used to integrate data on various aspects of cardiac performance, obtained from a variety of sources, in a standardized way. This approach provides an invaluable, highly detailed and dynamic map of the heart that clinicians can use to characterize a particular patient's function against the range of functional characteristics derived from large populations of patients, with the objective of allowing more precise evaluation of disease and targeting of therapies.

Computational modeling techniques are already being applied in various biomedical projects around the globe, including the Physiome Project [1] which describes whole-body physiology, the International Consortium for Brain Mapping (ICBM) [2], Informatics for Integrating Biology and the Bedside (i2b2) [3], and the Integrative Biology Project [4], to name a few. The Center for Computational Biology (CCB) [5] at UCLA, which hosts the ICBM, provides a number of infrastructural and middleware tools, mainly in the area of brain mapping. In the cardiac domain, the Cardiovascular Research Grid (CVRG) at Johns Hopkins University [6] provides grid computing infrastructure for cardiac research, inspired by the Cancer Bioinformatic Grid (caBIG) [7] and the Biomedical Informatics Research network (BIRN) [8].

The Cardiac Atlas Project (CAP) is a NIH sponsored, international, multi-institutional endeavor which aims to facilitate large scale statistical analysis of heart shape, structure, function and wall motion characteristics across various population groups, using parametric mathematical modeling tools. The initial goals of the Project are i) to develop a database of cardiac Magnetic Resonance Images (MRI) and associated patient data, ii) to develop standardized procedures for the contribution, curation, archival, classification, and sharing of data and derived analyses, and iii) to provide open source software for the mapping and analysis of cardiac morphometry, with particular emphasis on the spatio-temporal characteristics of regional heart wall motion. In collaboration with the CCB, infrastructure created for brain mapping research is being translated to the cardiac domain. The CAP is also developing software tools for accessing and analyzing cardiovascular imaging data, as well as procedures and policies for secure, ethical and efficient data and resource sharing. This paper provides an overview of the design and goals of the CAP, and describes the complex administrative and procedural issues related to the contribution of data to the CAP, standardization and sharing of data and software tools, and the protection of the rights of participants, contributors and users of the database.

2 Data

The CAP database currently includes image and text data of approximately 3000 subjects from two main early contributing studies: the Multi Ethnic Study of Atherosclerosis (MESA) [9] and the Defibrillators to Reduce Risk by Magnetic Resonance Imaging Evaluation (DETERMINE) [10] clinical trial. Several other smaller research studies are also contributing data.

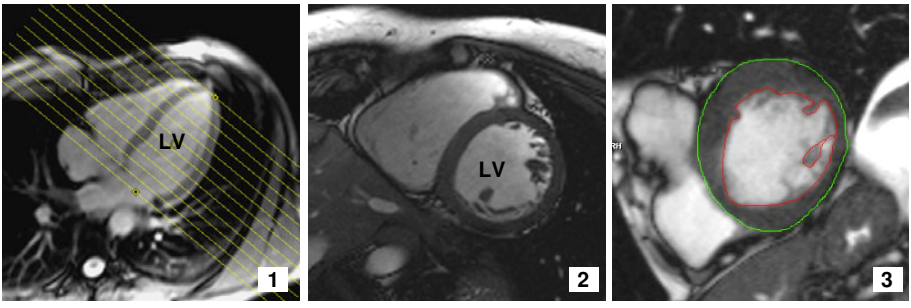

Fig. 1. Routine Cardiac MR images: 1) Cine image of long-axis view of the heart (LV = Left Ventricle) showing planned short-axis imaging planes (yellow lines); 2) Mid ventricular short-axis image; 3) Image analysis contours drawn around the inner (red) and outer (green) borders of the LV wall.

Individual datasets comprise cardiac MR images in DICOM format, together with image analysis files in the form of contours already drawn around the inner and outer borders of the left ventricle by the contributing study (Figure 1). Cases from all studies include cine image series acquired in the short and long axis planes of the heart. The DETERMINE cases also include delayed enhancement viability MR images used for detection and quantification of myocardial infarct.

Text data from the image DICOM headers (e.g. MRI pulse sequence type, image position and orientation, and other MR scan parameters) are automatically extracted and stored in the CAP database. Some limited clinical information is also contributed, including: age (years), gender (M/F), height (cm), weight (kg), systolic and diastolic blood pressure (mmHg), hypertension (y/n), heart rate (bpm), race/ethnicity (class), and classifications for hypertension, diabetes, smoking (Y/N), alcohol (Y/N), angina (y/n), ECG and NYHA classification.

These data are currently stored in two main databases within the CAP: a production database hosted by the CCB at UCLA, and a research database hosted by the University of Auckland, New Zealand.

3 Regulatory and IRB Requirements

Only data that were originally acquired with the approval of a local Institutional Review Board or Ethics Committee, and with informed consent from the participant compatible with data-sharing, may be contributed to the CAP. In observance of the USA HIPAA (Health Insurance Portability and Accountability Act) laws which protect participants' private health information, data must be de-identified at source before upload to the CAP database. Any data that can be used to identify an individual, e.g. names, dates (except for year), social security or medical record numbers, locations or device identifiers, are deleted. At no stage is any identifiable data read, analyzed or stored on any CAP computers.

Some contributing studies, such as DETERMINE, include a section in their own participant information sheet and consent form on the contribution of de-identified

data into the CAP. Participants can give or withhold consent to contribute their data to CAP independent of participation in the contributing study.

Other studies, such as MESA, are inherently designed for data-sharing and have an IRB approved informed consent process compatible with contribution to de-identification and sharing in the CAP project. In the case of MESA, explicit amendments were obtained from each field center's local IRB for the contribution of image and text data to CAP. Only data from those field centers with IRB approval, and those participants with informed consent compatible with CAP, were included in the CAP database.

The Cardiac Atlas Project Investigators themselves obtained the necessary IRB and Ethics Committee approvals to undertake the project at the two CAP centers- the University of Auckland, New Zealand, and the University of California Los Angeles, USA.

4 CAP Policies and Procedures

In order to ensure that all data provided to CAP are managed according to well defined principles, in accordance with the regulatory and ethical requirements associated with de-identified human image and clinical data, a number of policies and procedures related to data ownership, control and sharing have been developed. These policies apply to participants from whom the data is obtained, contributing studies which originally collected and have contributed the data, the CAP investigators and third-party Users who wish to access CAP data. The flow of data in CAP is shown in Figure 2.

4.1 Participants

Participants consent to contribute their de-identified image and text data to the heart disease research community now and in the future. All data are de-identified in a manner compatible with the HIPAA privacy rule, using the CCB's de-identification software with customized mappings [11] where original study identifiers, and private health information are replaced by CAP codes. This occurs at the site of the Contributing Study before upload to the CAP data servers, so the CAP does not receive or retain the original identifiers. There is, therefore, no possibility that CAP investigators, or third party Users of CAP data, can identify individuals, and all researchers must agree not to attempt to identify participants. The key linking CAP codes with original identifiers is retained by the Contributing Study, so that investigators of the Contributing Study could link results from CAP back to the original study if desired. Participants can request withdrawal of their data from the database at any time by requesting removal either via the CAP or directly to the Contributing Study. In this case the Contributing Study must notify CAP of which CAP codes must be deleted.

4.2 Contributors

Each Contributing Study has made substantial monetary, intellectual, and time investments for the collection of the data in a well-controlled manner (*viz.* original study design, recruitment, quality control, analysis, etc.), which represents a valuable scientific resource. Data contributed to CAP is therefore considered the property of

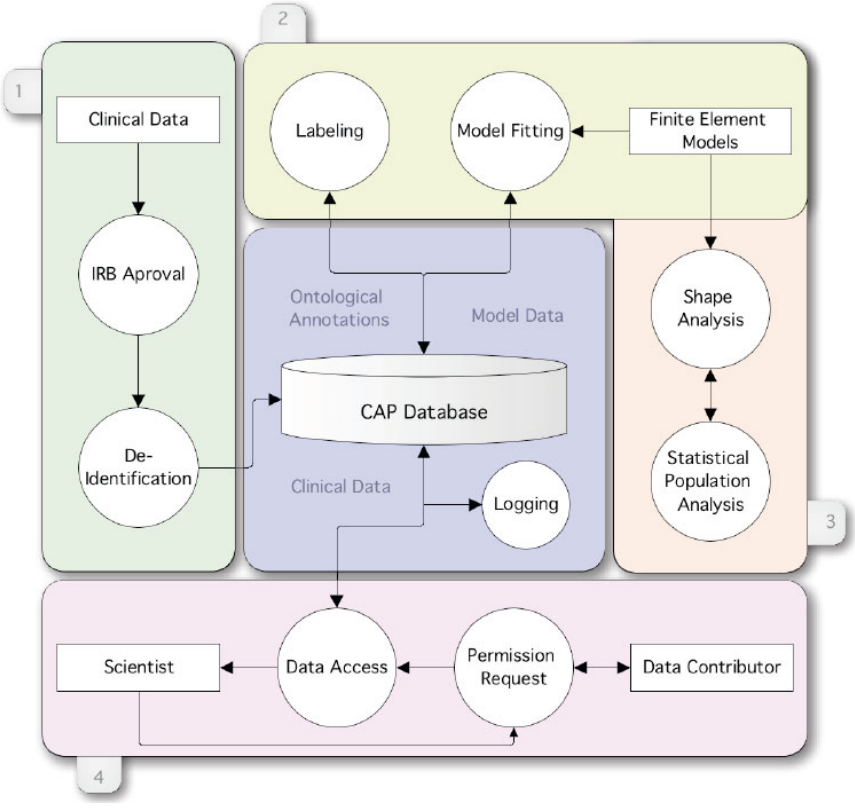

Fig. 2. CAP work-flow diagram visualizing 1) Data Acquisition; 2) Data Processing; 3) Data Analysis and 4) Public Data Access.

the Contributing Study. The Contributing Study Steering Committee controls all use of their data through data distribution agreements, on a case by case basis, as described below.

4.3 Users

Potential CAP users are required to submit a Research Proposal, outlining the rationale and goals of the project, term, and data storage, to the CAP Steering Committee. If acceptable, CAP will liaise with each of the Contributing Studies whose data is required for the Research Project. Each Contributing Study (or nominee) will then review the proposal and assess its eligibility with respect to the goals of the Contributing Study. If the proposal is approved, the User will be required to sign and abide by a Data Distribution (DDA) agreement for each of the Contributing Studies involved. Separate DDA's are required because terms and conditions governing data use are specific to the goals and rationale of each Contributing Study. The DDA defines terms and conditions of the use of the data, including publication policy, acknowledgements, security and intellectual property.

4.4 Intellectual Property

All software produced by the Cardiac Atlas Project (CAP) is freely available, via the CAP website, to researchers and educators in the non-profit sector, such as educational institutions, research institutes, and government laboratories.

CAP database and heart modeling tools, comprising database management, uploading and downloading of images, web browser interface, conversion of data formats, visualization, and parametric modeling of shape and motion, developed using CMGUI (the open source finite element modeling package developed by the University of Auckland Bioengineering Institute) [12], are being made available using the Mozilla Public License (MPL).

Commercialization of enhanced or customized versions of the software, or incorporation of the software or pieces of it into other software packages, is permitted subject to the terms of the license. Researchers are permitted to modify the source code and are strongly encouraged to share modifications with other researchers as well as with the CAP. Intellectual Property pertaining to the endpoints or specific aims of a Contributing Study, for example evaluation of a therapeutic drug or device which formed the primary hypothesis of the Contributing Study, will in general remain the property of the Contributing Study.

Intellectual property developed by third party researchers using CAP data, not relating to the specific character of the Contributing Studies, should remain the property of the developers.

5 End User Tools

5.1 Database

The CAP Database (Figure 3), hosted by the CCB at UCLA, builds upon existing brain mapping infrastructure, which has been modified for cardiac images. Access to the database is secure and privileges are assigned based on the User's needs and intent.

Browsing: The user will be able to browse the data in the CAP database. Data will be sortable by a few key fields, such as Research Group (e.g. DETERMINE), imaging protocol series description (e.g. TruFISP), age, etc.

Searching: The user will be able to perform simple and advanced queries in order to search for data. A simple query might be based on two or three fields (e.g. male + MESA + diabetic) or may contain a 'wildcard' field. (e.g. diabetic). An advanced query would allow the user to search for cases more specifically (e.g. male + MESA + diabetic + 25 < age < 45 years + acquisition date > 2003).

Downloading: Once the User has decided which data s/he would like to use, s/he would need to submit a Research Proposal to CAP for approval. Data may be downloaded only upon the execution of a completed Data Distribution Agreement.

5.2 CAP Client

MRI data may be visualized and patient specific mathematical models created using the open source (Mozilla tri-license) CAP client [13] which can be freely downloaded at the Project website.

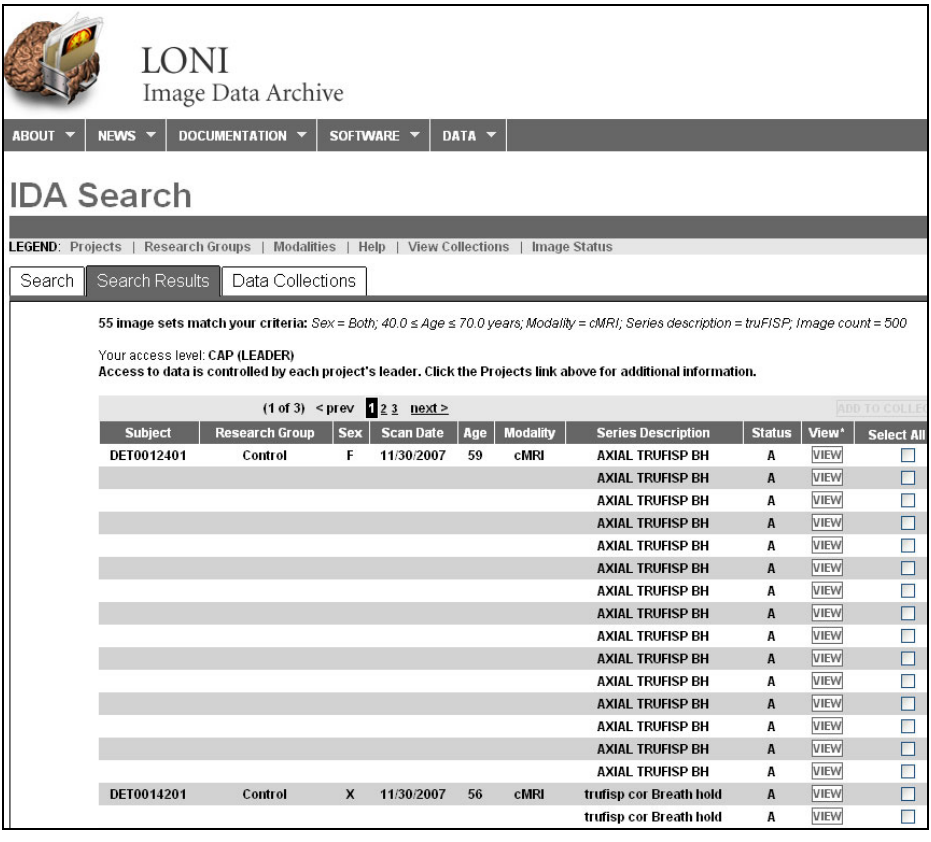

Fig. 3. Example of Search results for the CAP Database hosted by the CCB at the UCLA Laboratory of Neuro Imaging (LONI)

The CAP client provides:

- Retrieval of MRI data from the database across the network and the capability to upload models generated from such data to the database.
- Visualization of 2D CMR images, 3D visualization of the mathematical model constructed from the CMR images, and 4D Visualization (3D Visualization with the temporal dimension as the 4th dimension).
- Fitting of a mathematical model to a series of CMR images with minimal human intervention, enabling a large set of data to be automatically fitted to appropriate models and parameters. The Client software will also provide a means for the human user to be able to interactively and graphically modify the relevant model parameters.
- Tools necessary for the statistical analysis of the data. This will be used for the generation of the parametric distribution models.

The CAP Client is developed in C++ and runs on the Windows, Mac OS X and Linux platforms.

6 Data Analysis

Atlas based methods are now well established for the statistical classification and quantification of shape and wall motion characteristics [14]. These methods enable standardized analysis of statistical variations present within and among patient groups, and enable classification of individual phenotypes within known population distributions. In almost all cases contributed to the CAP, contours are contributed in association with the images and clinical information. These contours can be used in a standardized model-based analysis to establish shape and motion with respect to a standard coordinate system, similar to the Talairach coordinate system used in the brain. Since shape and motion are mathematically mapped, statistical tools such as principal component analysis can be used to quantify the significant modes of variation present in a population. In a preliminary analysis, major modes of variation within the DETERMINE cohort were associated with size, sphericity and mitral valve geometry, each of which are known indices of geometric remodeling. Projection of an individual's shape and motion onto these modes (e.g. sphericity) provides a standardized method for quantifying the amount of each mode present.

7 Future Directions

In accordance with the goals of standardized classification and sharing of data and resources, the CAP is developing and building upon currently available ontological schema to describe the data and will federate cardiovascular modeling software and data to make them available to the cardiovascular research community via the Cardiovascular Research Grid (CVRG) [6].

The CAP database will be interfaced with the CVRG-Core, and modified to implement interfaces and mechanisms compatible with CVRG enabled analysis tools. The CAP client software will also be grid-enabled, in order to be used in standard CVRG workflows, including a portal component to enable interaction with other resources on the grid. The parametric modeling tools and associated ontological schema that are being developed by CAP will be designed to facilitate data fusion between different imaging protocols and modalities as well as other data sources.

The standardized classification and description of CAP data elements (CV images and derived morphological information including contours and parametric geometry descriptions) will occur through registration of an information model and associated semantic annotations, expressed in the Web Ontology Language (OWL) [15], at the National Center for Biomedical Ontologies (NCBO) [16]. This will allow grid-enabled tools to query and access data of the correct type, and databases to declare what type of data are available. Data and derived results from several studies can be labeled and collated in a standardized manner to achieve meta or subgroup analyses via the use of existing domain ontologies, such as the Foundational Model of Anatomy (FMA) for anatomical data [17] and the Systematized Nomenclature of Medicine - Clinical Terms (SNOMED CT) for clinical terms [18]. Where gaps occur, suggested terms will be proposed based on feedback from the radiological and cardiological communities via online resources such as the NCBO BioPortal [16] and WebProtégé [19].

8 Conclusions

The CAP currently hosts approximately 3000 cardiac MRI studies, derived functional analyses and associated subject characteristics data which represents a substantial and valuable resource. Tools for the de-identification of data were developed and validated. These tools were provided to the Contributing studies and were used successfully. The necessary IRB and Ethics Committee approvals were obtained and policies were developed to protect the rights of subject participants, contributors and users of the database. Applications to use the data can now be submitted to the CAP. Applicants who are granted access can browse and query the database as well as view the images therein, and can download the data upon completion of a Data Distribution Agreement. The CAP client, which allows model generation, is also available for download at the Project's website.

Acknowledgement

The project described was supported by Award Number R01HL087773 from the National Heart, Lung, and Blood Institute. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Heart, Lung, And Blood Institute or the National Institutes of Health.

References

1. ABI Physiome Home Page, http://www.physiome.org.nz/
2. International Consortium for Brain Mapping, http://www.loni.ucla.edu/ICBM
3. Informatics for Integrating Biology and Bedside Home Page, http://www.i2b2.org/
4. Integrative Biology Project Home Page, http://www.integrativebiology.ox.ac.uk/
5. Center for Computational Biology Home Page, http://www.loni.ucla.edu/ccb
6. Cardiovascular Research Grid Home Page, http://www.cvrgrid.org/
7. Cancer Biomedical Informatics Grid Home Page, http://cabig.nci.nih.gov
8. Biomedical Informatics Resarch Network Home Page, http://www.nbirn.net/
9. Bild, D.E., Bluemke, D.A., Burke, G.L., Detrano, R., Diez Roux, A.V., Folsom, A.R., Greenland, P., Jacob Jr., D.R., Kronmal, R., Liu, K., Nelson, J.C., O'Leary, D., Saad, M.F., Shea, S., Szklo, M., Tracy, R.P.: Multi-ethnic study of atherosclerosis: objectives and design. Am. J. Epidemiol. 156, 871–881 (2002)
10. Kadish, A.H., Bello, D., Finn, J.P., Bonow, R.O., Schaechter, A., Subacius, H., Albert, C., Daubert, J.P., Fonseca, C.G., Goldberger, J.: Rationale and design for the Defibrillators to Reduce Risk by Magnetic Resonance Imaging Evaluation (DETERMINE) trial. J. Cardiovasc. Electro-physiol. 20, 982–987 (2009)
11. LONI De-identification Debabelet, http://www.loni.ucla.edu/Software/DiD
12. CMGUI software, http://www.cmiss.org/cmgui
13. Backhaus, M., Britten, R., Chung, J.D., Cowan, B.R., Fonseca, C.G., Medrano-Gracia, P., Tao, W., Young, A.A.: The Cardiac Atlas Project: Development of a Framework Integrating Cardiac Images and Models. In: MICCAI 2010, Workshop on Statistical Atlases and Computational Models of the Heart: Mapping Structure and Function plus a Cardiac Electrophysiological Simulation Challenge (STACOM-CESC 2010), Beijing, China. LNCS (2010) (in press)

14. Young, A.A., Frangi, A.: Computational cardiac atlases: from patient to population and back. Experimental Physiology 94(5), 578 (2009)
15. OWL Web Ontology Language, http://www.w3.org/TR/owl-features/
16. Bioportal 2.0 Home Page, http://www.bioontology.org/tools/alpha.html
17. Foundational Model of Anatomy Home Page, http://sig.biostr.washington.edu/projects/fm/AboutFM.html
18. Ryan, A.: Towards semantic interoperability in healthcare: ontology mapping from SNOMED-CT to HL7 version 3. Australian Computer Society, Inc., Hobart (2006)
19. WebProtégé Home page, http://protegewiki.stanford.edu/index.php/WebProtege