31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Computing Human Image Annotation

David S. Channin MD, Pattanasak Mongkolwat PhD, Vladimir Kleper, and Daniel L. Rubin MD MS

Abstract— **An image annotation is the explanatory or descriptive information about the pixel data of an image that is generated by a human (or machine) observer. An image markup is the graphical symbols placed over the image to depict an annotation. In the majority of current, clinical and research imaging practice, markup is captured in proprietary formats and annotations are referenced only in free text radiology reports. This makes these annotations difficult to query, retrieve and compute upon, hampering their integration into other data mining and analysis efforts.**

This paper describes the National Cancer Institute's Cancer Biomedical Informatics Grid's (caBIG) Annotation and Image Markup (AIM) project, focusing on how to use AIM to query for annotations. The AIM project delivers an information model for image annotation and markup. The model uses controlled terminologies for important concepts. All of the classes and attributes of the model have been harmonized with the other models and common data elements in use at the National Cancer Institute. The project also delivers XML schemata necessary to instantiate AIMs in XML as well as a software application for translating AIM XML into DICOM S/R and HL7 CDA. Large collections of AIM annotations can be built and then queried as Grid or Web services.

Using the tools of the AIM project, image annotations and their markup can be captured and stored in human and machine readable formats. This enables the inclusion of human image observation and inference as part of larger data mining and analysis activities.

I. Introduction

Clinical and research use of medical imaging as a biomarker or endpoint in image based clinical trials requires identification of lesions, their characterization and their measurement in a fashion that is structured, computable, and semantically interoperable. Most clinical and research medical imaging is now done with digital imaging technology. The vast majority of these digital medical images are stored and manipulated as Digital Imaging and Communications in Medicine (DICOM) standard [1] objects. DICOM contains a large amount of metadata about whom, how and when the images were acquired. It also specifies precisely how the information is stored in the image object and how pixels should be interpreted for display. But, DICOM says nothing about the meaning of the pixel information; the extraction of which is usually a task for a trained human observer (though in some specific, niche cases, there are machine observers). Furthermore, DICOM has no defined, standard mechanism for representing these observations and the markup that typically accompanies them.

Imagine a large picture archiving and communication system (PACS) at a clinical or research institution. There are millions of studies and hundreds of millions of images. It is relatively straightforward, though possibly non-trivial in any given commercial environment, to query this collection to, "find all the CT studies of the chest with a slice thickness less than or equal to 1.5 mm". Performing, however, the query, "find all the CT studies of the chest with a slice thickness less than or equal to 1.5 mm that have a spiculated nodule greater that 1cubic centimeter in the right upper lobe of the lung", is currently impossible. The truly useful cancer research queries, "find all the CT studies of the chest with a slice thickness less than or equal to 1.5 mm that have a spiculated nodule greater that 1cubic centimeter in the right upper lobe of the lung that has decreased in volume by 20%" are beyond consideration in today's imaging environment.

The mission of the National Institutes of Health's (NIH) National Cancer Institute's (NCI) Cancer Bioinformatics Grid (caBIG™) [2] is to provide infrastructure for creating, communicating and sharing bioinformatics tools, data and research results, using shared data standards and shared data models across the funded Comprehensive Cancer Centers and the cancer research community in general. Recognizing the gap in the ability to capture, store, query, retrieve and manipulate image annotations and markups, the NCI funded the Annotation and Image Markup (AIM) project.

The AIM project defines an ontology of annotation and image markup, a UML information model and provides the extensible markup language (XML) artifacts for creating them. A software toolkit and a reference software development kit (SDK) and a validation and transcoding tool allow for serialization of the AIM as native XML, DICOM structured reporting (DICOM S/R) [3] or HL7 Clinical Document Architecture (HL7 CDA) objects [4].

The long term vision is for large collections of annotations to be created in conjunction with the already large collections of clinical and research medical images. This will allow query of annotation, not only for retrieval of relevant images, but for correlation of image observations and their characteristics (as a proxy for phenotypic expression) with other biomedical data including genomic expression.

Manuscript received April 7, 2009. This work was supported, in part, as a task order under contract 79596CBS10 as part of the National Cancer Institute's caBIG™ program to The Robert H. Lurie Comprehensive Cancer Center of Northwestern University.

D. S. Channin is with the Department of Radiology, Northwestern University Feinberg School of Medicine, Chicago, IL 60611 USA (phone: 312-926-9165; fax: 312-926-5991; e-mail: dsc@northwestern.edu) .

P. Mongkolwat is with the Department of Radiology, Northwestern University Feinberg School of Medicine, Chicago, IL 60611 USA (e-mail: p-mongkolwat@northwestern.edu).

V. Kleper is with the Department of Radiology, Northwestern University Feinberg School of Medicine, Chicago, IL 60611 USA (phone: 312-926-9165; fax: 312-926-5991; e-mail: vkleper@northwestern.edu).

D. L. Rubin is with the Department of Radiology and the Center for Biomedical Informatics Research, Stanford University, Stanford, CA 94305 (e-mail: dlrubin@stanford.edu).

II. Primary Use Case

The primary use case is the setting of an image based clinical trial. There is a DICOM series of images from one patient at a first point in time and a second DICOM series of images from the same patient at a second time point. Observer 1 names, annotates and measures lesions in the first series of images. For each named lesion in the first series of images, observer 2, possibly different from observer 1, possibly blinded to observer 1's results, annotates and measures the same lesions. A researcher, perhaps an adjudicator in a clinical trial, accesses both series of images, the AIM from observer 1 and the AIM from observer 2. The researcher may perform calculations on the two sets of measurements. The researcher then creates an annotation that binds the individual image annotations together.

This use case is broad enough to handle any number of clinical or research annotation scenarios. It devolves rapidly to the simple clinical or teaching annotation use case.

III. The AIM Model

The AIM model consists of over 40 classes. The model adheres to the NCI caBIG's silver compatibility guidelines [5] of controlled vocabularies, common data elements (CDEs), well-documented APIs and UML models. The AIM version 1 Rev 12 model has passed caBIG silver compatibility review.

The AIM model provides for two types of primary annotations: ImageAnnotation and AnnotationOfAnnotation, both subclasses of the Annotation class. As described above, Image Annotation is used to annotate a single lesion in a single series of images at a single point in time. AnnotationOfAnnotation is used to bind ImageAnnotations together often via calculations and characterizations between them.

The model has classes for patient, user and equipment information. It contains a set of classes related to geometric shape and the coordinates used to represent them as well as classes for referencing DICOM and web (URL) images. The model also contains classes for calculations and calculation results.

The classes of primary interest, however, are those for ImagingObservation, ImagingObservationCharacteristic and AnatomicEntity. In the version 2 model, now under development, the AIM model will add classes for Inference, AnatomicEntityCharacteristic, and Ratings for both ImagingObservationCharateristic and AnatomicEntityCharacteristic. These classes are the key to semantically interoperable annotations. Each requires the use of terms from controlled vocabularies such as RadLex [6],[7], SNOMED-CT [8], or UMLS [9]. It does this by providing attributes to indicate the coding scheme designator (e.g., RadLex), the coding scheme value of the term in question, and the coding scheme meaning of the term.

The complete details of the AIM model, including how to download the model, schemata, software and documentation are described elsewhere [10].

IV. Worked Example

Consider a computed tomography study of the chest for a pulmonary nodule identified on prior chest radiograph. The DICOM study consists of a series of contrast enhanced CT images of the chest. Suppose that the nodule appears on 5 of these images each with a unique DICOM Sop instance Universal Identifier (UID). We create an ImageAnnotation with the *type* "Target_Lesion". Note that the *type* attribute of the ImageAnnotation class is (now) an NCI Common Data Element (CDE). As such, it has a precise definition and, in this case, a value domain. In version 2.0 of the model, the complete value domain for ImageAnnotation *type* is, "New_Lesion", "Clinical_Finding", "Non_target_lesion", "Non_lesion_at_baseline", "Quality_control", "Research", "Target_Lesion", and "Teaching". These are the types of ImageAnnotations that are currently allowed. Note that there is a well defined process to expand the value domain of an NCI CDE should that be necessary.

Our example annotation will have other attributes of annotations, date of creation, a set of user and equipment attributes and a set of DICOM classes to reference the DICOM images. Note that the NCI uses a different model for the header information in DICOM images and AIM just references the DICOM objects.

In this example, the primary AnatomicEntity attributes might be the tuple, (RadLex, RID1303, "upper lobe of right lung"). Similarly, for ImagingObservation we might assign, (RadLex, RID3875, "nodule"). We might then assign the ImagingObservationCharacteristic of (RadLex, RID5713, "spiculated margin") and so on.

For markup, depending on the software application we are using, we may have a set of polygons around the nodule in each of the 5 image planes. These would be captured as 5 geometric shapes each with a set of two dimensional spatial coordinates. The model also supports three dimensional spatial coordinates.

When we make a calculation on this lesion, we choose a calculation *type*, again, from the value domain of the calculation type CDE. In this case, we select "Volume_estimated_from_three_or_more_non_coplanar_2D_regions". The result of the calculation is stored in a set of classes related to calculation results. Calculation results can be scalar, vector, histogram, matrix or array.

V. Querying

Once AIM annotations, as described above, are created, they can be stored in a local database and queried and retrieved as necessary. Such a use has already been described [11] and presents a major advantage of AIM in enabling query access to the semantic content in images.

In addition to querying AIM stored in a local database, it is possible to query AIM via caGrid query. In caBIG, a silver compliant model, such as AIM, can be used to create grid data services. In fact, the caCORE SDK [12],[13]

contains many tools to automate the process of creating such a service. These grid services can receive and store AIM instances and respond to queries for them. One example of these services is the National Cancer Image Archive (NCIA) which contains a large number of collections of research images from various projects [14]. The software for NCIA itself is free and open source and so it is easy to imagine other instances of NCIA running on the Cancer Grid (caGRID) and in fact, there are. NCIA has a grid facing service that serves the NCIA Image model, a subset of the DICOM header information. Very shortly, it will have a grid service serving the AIM model.

NCI has recently funded an image query formulation project to exploit the AIM and NCIA Image models. In this project, the AIM ontology is used to drive a user interface, a mock-up of which is shown in Fig. 1.

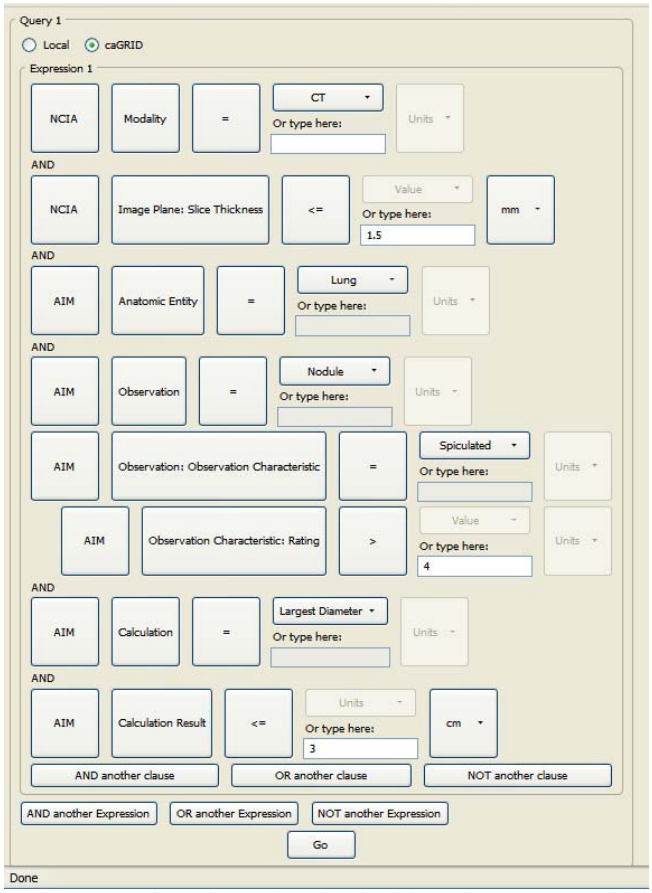

Fig. 1. A mock-up of a user interface for federated query across multiple caBIG information models.

The user selects the model and the attributes of interest for the query. The AIM ontology constrains choices for predicates and attributes to the appropriate value domains.

The application can then create a canonical query graph that represents the selection of the user. This query graph is then translated into DCQL, the Distributed Common Query Language [15], an extension of CQL. "CQL (Common/caGrid Query Language) is the caGrid query language used for all caGrid Data Services to express queries against a data source using an object oriented language"[16].

> "DCQL, the language used to express federated queries, is an extension to CQL, the language used to express single data service queries. Both CQL and DCQL use a declarative approach to describe the desired data by identifying the nature of the instance data with respect to its containing UML information model. That is, a query can be seen as identifying a class in a UML model, and restricting its instances to those which meet criteria defined over that class's UML attributes and UML associations. The primary additions to CQL, which DCQL provides, are the introduction of the ability to specify multiple target services (aggregations), and the ability to specify object restrictions through relationships to Objects on remote data services (joins)."[15].

The same query graph could be translated into SPARQL [17] and we are investigating the feasibility of translating SPARQL into CQL/DCQL.

The DCQL query is passed to the caGRID Federated Query Processor (FQP) [18]. FQP is the mechanism to perform basic distributed aggregations and joins of queries over multiple data services.

Each of the AIM and NCI Image model data services process the query and return instances that satisfy the queried attributes of the (respective) models. The querying application can then use DICOM (and other) UIDs to match annotations and images for use in further processing.

VI. Conclusion

Image annotations capture the essential meaning of pixel information from human observers. Characterizing lesions, measuring them and their growth are critical to understanding pathological processes. Perhaps more importantly, correlating these imaging observations with other observations and data about the patient will be core to our complete understanding of disease.

The National Cancer Institute's caBIG project has developed tools for performing these kinds of analysis and correlation across different models of biomedical information. Recent work has extended this model of collaboration to images and now, to image annotation.

Widespread adoption of the AIM model and its representation in both DICOM and HL7 standard formats will advance our ability to document and understand disease phenotype through the proxy of biomedical imaging.

References

[1] Digital Imaging and Communications in Medicine (DICOM) Standard. PS 3 2008. Washington DC: National Electrical Manufacturer's Association (NEMA). http://medical.nema.org/dicom/2008/. Accessed April 3, 2009.

[2] Von Eschenbach AC, Buetow K. Cancer Informatics Vision: caBIG™. Cancer Informatics 2: 22-24.
[3] Clunie, D. A., DICOM Structured Reporting, PixelMed Publishing, 2000.
[4] Health Level 7. Ann Arbor Michigan. Health Level 7, Inc. http://www.hl7.org/ Accessed April 3, 2009.
[5] NIH NCI caBIG Compatibility Guidelines [Online]. Available: https://cabig.nci.nih.gov/guidelines_documentation/compat_v3/. Accessed April 3, 2009.
[6] Langlotz CP. RadLex: a new method for indexing online educational materials. Radiographics. 2006 Nov-Dec; 26(6):1595-7.
[7] RadLex. http://www.radlex.org/. Accessed April 3, 2009.
[8] SNOMED-CT. International Health Standards Development Organization. http://www.ihtsdo.org/snomed-ct/. Accessed April 3, 2009.
[9] Unified Medical Language System. National Institutes of Health National Library of Medicine. http://www.nlm.nih.gov/research/umls/. Accessed April 3, 2009.
[10] AIM JDI
[11] Rubin DL, Rodriguez C, Shah P, Beaulieu C. iPad: Semantic Annotation and Markup of Radiological Images. AMIA Annu Symp Proc. 2008 Nov 6:626-30.
[12] Phillips J, Chilukuri R, Fragoso G, Warzel D, Covitz PA. The caCORE Software Development Kit: streamlining construction of interoperable biomedical information services. BMC Med Inform Decis Mak. 2006 Jan 6; 6:2.
[13] NIH NCI Cancer Common Ontologic Representation Environment (caCORE). https://cabig.nci.nih.gov/index_html/concepts/caCORE_overview?pid=primary.2006-07-07.4911641845&sid=caCORE_Overview&status=True .Accessed April 3, 2009.
[14] NIH NCI The National Cancer Imaging Archive. http://ncia.nci.nih.gov/. Accessed April 3, 2009.
[15] NIH NCI caBIG Distributed Common Query Language. http://cagrid.org/display/fqp/DCQL. Accessed April 3, 2009.
[16] NIH NCI caBIG Common Query Language. http://cagrid.org/display/dataservices/CQL. Accessed April 3, 2009.
[17] Prud'hommeaux E, Seaborne A., Eds. SPARQL Query Language for RDF,W3C Recommendation, 15 January 2008, http://www.w3.org/TR/rdf-sparql-query/ . Accessed April 3, 2009.
[18] NIH NCI Federated Query Processor. http://cagrid.org/display/fqp/Home . Accessed April 3, 2009.