

Mayo CTSA Symposium
Standards for Translational-Research Interoperability
Friday, September 28th,2007

Summary

Standards for Translational Research Interoperability Rational, Use Cases, IRB, SDO Efforts; CG Chute MD DrPH, Mayo Clinic

An overview of why translational research would benefit from data interoperability, and how that interoperability in turn depends upon comparable and consistent data was proposed. The role of data standards as a mechanism to achieve comparable and consistent data was explored, together with some review of major standards efforts including HITSP (Health Information Technology Standards Panel), HL7, and related activities.

Clinical Data Interchange Standards Consortium (CDISC); Ed Helton PhD, Chair Elect, CDISC

The CDISC body of data standards, and their application to clinical research was reviewed. In particular, the BRIDG model for clinical trials data, as developed by CDISC in partnership with NCI's caBIG (Cancer Biomedical Informatics Grid), FDA, and HL7, provided the best example of cross-SDO (Standards Development Organization) cooperation on a problem highly-relevant to the CTSA community.

OCRe: The Ontology of Clinical Research; Ida Sim, MD PhD, USCF Center for Clinical Informatics

The domain of clinical trials has been the subject of standards by BRIDG, caBIG, HL7, and many standards organization. Dr. Sim described a collaboratively developed ontology that may form the descriptive backbone for many of these standards, in partnership with Alan Rector at the University of Manchester, UK, the UK Cancer Grid, and the NCI's caBIG. This ontology has clear applicability to organize CTSA research.

caBIG and Its Potential Starting Points; George Komatsoulis, PhD, NCI Center for Bioinformatics

Given the broad recognition that the maturing NCI caBIG program may serve as a practical source of standards, starting points, and organization model for CTSA, Dr. Komatsoulis outlined the Program structure and sociology within caBIG including governance, data sharing, and open development/open source. Additionally, a detailed description followed of the technology basis, specifically the requirements for compatibility, modeling in a federated environment, and the manifestation of a shared technology platform using caGrid.

NLP Collaborative Standards for Patient Record Research; Guergana Savova, PhD, Mayo Clinic

Translational research requires a high degree of data granularity and detail beyond that available in most clinical data repositories, registries, or administrative databases. One mechanism to collect this information is by the parsing of clinical texts such as summary notes, discharge summaries, radiology reports, and procedure notes. Dr. Savova overview the state of the art in this domain, including the UIMA (Unstructured

Information Management Architecture) open-source resource for assembling and executing NLP (natural language processing) annotators.

Discovery Workflow in Bioinformatics Analyses, Brian Wilson, Mayo Clinic

The analyses of complex bioinformatics data entails sequential steps of quality evaluation, normalization, inferencing, and interpretation. Some of these steps are monotonous and tedious, some require careful investigator attention. Many techniques of automating workflow templates, to standardize and simplify such analyses have emerged and were described. Furthermore, editing and sharing such workflows has become a reality that can leverage all translational researchers in a scalable fashion.

Discussion

A full two hours of discussion explored how these standards are relevant, how they can be adapted and shared, and how they should evolve. A major consensus was the balance between simple IT infrastructure and informatics science that can advance the art – both are necessary. The spectrum of work was universally considered large.

Action Items

To provide a practical focus with short-term deliverables, the group adopted an initiative by Ida Sim to modify clinical trial registry infrastructure to build a standards-based description of ongoing human studies research across the CTSA consortium. This as the advantage of providing a useful catalog of activities to investigators and NIH, while at the same time exercising many disparate vocabulary and information model standards to capture and display this information.

Drs. Sim and Chute were charged to outline a more detailed proposal for consideration by the CTSA informatics group [done].

Engagement of the new “class” of CTSA members is a high priority. Dissemination of the symposium resources as video and slides on a public web site was assigned to Dr. Chute [done].

Availing the AMIA meeting opportunity to assemble CTSA investigators to pursue further standards adoption was proposed. [scheduled].