Ontology Boot Camp Jan 26th - 27th 2006

The first of several Ontology Boot Camps was hosted at the BIRN Coordinating Center at UCSD's School of Medicine last week. Thirteen scientists from UC Irvine, UCLA, Stanford, UCSD, Duke, University of Tennessee Memphis, University of Texas Health Science Center San Antonio, the NIH, and Drexel University participated in person or virtually with more than a dozen scientists and programmers from UCSD. BIRN's Ontology mapping software, Bonfire, was an integral part of the workshop. The overarching workshop goal: Make BIRN databases and data understandable to human and machine!

The Ontology Task Force (Carol Bean and Maryann Martone, co chairs; Bill Bug, Jeff Grethe, Amarnath Gupta, Christine Fennema-Notestine, Jessica Turner) organized the two day workshop. Maryann offered inspirational words to the group during the opening sessions: "This [mapping to an ontology] is not an easy process. If it were, we would have this done. I don't think we are going to get it perfectly, but if we refine it, it will help."

Specific work shop goals were laid out:

1. Map databases to BIRN knowledge sources (Bonfire) and provide this mapping to BIRN mediator (Table names?; Field names?; Data values?).
2. Extend knowledge bases where necessary to accommodate BIRN data. Add concepts and link to existing concepts. Identify areas of difficulty and confusion and try to come to consensus.
3. Identify the set of ontologies currently used by BIRN and add them to Bonfire, e.g., NCBI taxonomy
4. Identify areas that are poorly covered where additional ontologies may be needed
5. Develop "Policies and Procedures" for working with ontologies in BIRN

The two day workshop was a surprising success, given the complexity of implementing consensus-derived solutions. Seven highlights of the meeting's results:

1) Function/Morph BIRN made progress on mapping terms to UMLS. They identified key concerns and limitations in using some of the UMLS concepts for generic human experimental descriptions, and by trial and error defined some "best practices" for submitting new concepts. They anticipate having the database partially annotated in the next six weeks so that they can demonstrate a semantic query via the mediator by the time of the FBIRN All Hands Meeting in March. Follow-up meetings with BrainMap collaborators regarding cognitive terms and concepts have been scheduled.
2) The Human Imaging Database table names and fields were identified exhaustively and the task of defining each one was begun. Function/Morph BIRN have divided up the task of mapping the remaining schema and contents. Follow-up between Morph BIRN and FBIRN HID experts and domain experts regarding annotation progress will occur in two weeks.
3) Members of the Mouse BIRN Ontology group made progress on mapping anatomical terms from the mouse atlas to the structural hierarchy provided by Neuronames. Mouse BIRN established a standard schema for entering species, strain, genotype and phenotype information on subjects.
4) All terms mapped by the sources to UMLS will be provided on the BIRN ontology page as "BIRNLex". Jessica Turner has already sent the list of terms form Human BIRNs.
5) Maryann and Jeff will assemble all of the newly mapped or added terms during the workshop and post to BIRN's ontology page. This list will serve as the basis for developing curatorial procedures for BIRN.
6) Policies and procedures were being formalized:
    a. The reuse and extension of existing ontologies is preferable.
    b. If two or more ontologies are substantially overlapping, one should be chosen for addition to Bonfire. Mapping between multiple ontologies may be desirable.
    c. Never use a concept or relation if it is not optimal.
    d. When adding and linking concepts, it is helpful to consider how someone would search for this data as a guide to what and how many relationships should be added. Don't worry about providing an exhaustive set of links, however.
    e. Always check the semantic type, definitions and existing links/synonyms when trying to select a concept from Bonfire.
    f. New concepts proposed for Bonfire MUST have a definition and a semantic type listed.
7) Users of the Bonfire browser could add terms and relationships. The members of the workshop suggested several improvements to Bonfire:
    a. Faster, (unanimous)
    b. Add "Source Vocabulary" to the display page.
    c. Allow search by semantic type and user name on the search and display pages.
    d. Improve the ability to select terms for the two node searches
    e. Remove some vocabularies that are not useful to BIRN, e.g., CSP.

f. Ability to edit entries after they are originally entered.
g. At least part of the Bonfire should be ported to Protégé.
h. Modify the entry tool to reject entries that do not have a definition and a semantic type.

At the beginning of the workshop, Bill Bug of Drexel University, who is credited with the inspiration behind the workshop, quipped, “This is the first meeting, so even the process is a work-in-progress.” The significant outcomes of the meeting, given this starting point, attest to the dedication and hard work by all of the boot camp participants.