

BioPAX Work Group
June 1, 2005 Conference Call Minutes

Participants: Michael Cary, Ken Fukuda, Elizabeth Glass, Gopal Gopinath, James Long, Joanne Luciano, Jonathan Rees, Mustafa Syed, Vincent Yau, Jeremy Zucker

Action Items:

- Mike – fix broken links in OWL-Docs, send announcement about documentation availability to biopax-discuss; request additional reviewers of documentation
- Joanne, Jonathan, Jeremy (and anyone else willing to do so) – review BioPAX Level 2 documentation, provide feedback to Mike and Gary
- Joanne – contact Frank Gibbons, Ian Donaldson, Robert Stevens and Danieli Turi regarding PSI-MI → BioPAX converter effort.
- Mike and Gary – set up a wiki page recommending standard practice by data providers in biopax data organization and access
- Mike – follow up with NIGMS to determine status of conference grant
- Mustafa, Jeremy and Joanne - set up an on-line hackathon session to complete the WIT conversion.
- Mike – set up web page for November meeting in Japan
- Jeremy – contact Peter about collaborating on the BioWarehouse Converter
- Ken – check on the availability of the facilities for a pre- and post-Face-2-face hackathon in Japan.

Summary:

1. Development status update
 - a. Gary and Mike finished the Level 2 documentation
 - i. Posted on the website (linked to from homepage)
 - ii. Documentation includes:
 1. 63 pages of text describing all features of BioPAX Level 2 (not just the new features since Level 1)
 - a. Describes classes and slots, best practices, sample use cases
 - b. Posted here:
<http://www.biopax.org/Downloads/Level2v0.9/biopax-level2-documentation.pdf>
 2. Updated class and slot definitions in the biopax-level2.owl file
 3. Auto-generated (from biopax-level2.owl) OWL-Docs
 - a. HTML pages generated by a Protégé plug-in
 - b. Some bugs still (broken links), Mike will fix and then send out announcement to biopax-discuss list
 - c. Can be accessed here:
<http://www.biopax.org/Downloads/Level2v0.9/OWLDoc/>

- iii. Documentation reviewers – we need people to read and comment on the documentation (primarily the 63 page doc, since the others are derivatives of it)
 - 1. Jeremy, Joanne, and Jonathan agreed to review documentation
 - 2. Other reviewers welcome
 - iv. Remaining steps to Level 2 ratification – what do we need to do before upgrading the beta release to the Level 2, version 1.0 final release (besides reviewing documentation)?
 - 1. Last step: Testing the Level 2 ontology
 - a. Required: Several small sample data sets
 - i. E.g. manual conversions of PSI-MI data
 - b. Desired: Largish data sets, several possibilities for these exist, such as:
 - i. Small set of detailed cancer pathways by Sander group and/or direct output from their cPath DB (which is mostly protein-protein interactions currently)
 - ii. PSI-MI → BioPAX converter
 - 1. Joanne will contact Robert Stevens, Danieli Turi, Frank Gibbons, Ian Donaldson to gauge their interest
 - v. Target date for Level 2 ratification remains June 25 (ISMB meeting)
2. LibBioPAX / conversion status
 - a. KEGG
 - i. Lemer group nearing completion of their KEGG → BioPAX converter
 - 1. They'll likely finish within a few weeks
 - 2. Christian Lemer sent several finished pathways to Gary and Mike, they looked very good
 - a. Each biopax file was a separate pathway map from KEGG mapped to a specific organism
 - b. WIT
 - i. Project nearing completion
 - 1. Jeremy, Joanne and Mustafa will set up an on-line hackathon session to address the remaining issues.
 - ii. Mustafa recently finished translating a set of membrane transport pathways, which he sent to Jeremy for review and comment
 - 1. Jeremy evaluated first version and sent back comments, he is currently waiting for corrected membrane transport reactions from Mustafa
 - c. BioWarehouse conversion
 - i. Jeremy asked if anyone knew of a tool that could store RDF in a relational DB
 - 1. Jim mentioned Castor, which can convert XML to Java objects
 - 2. Mike thought he recalled seeing other tools for this

- ii. The reason Jeremy asked was because he's trying to determine the most appropriate way to map between BioWarehouse and BioPAX and is interested in exploring the following possibility:
 - 1. If you could go back and forth from a relational DB into RDF, you could first use SQL queries to rearrange data in a relational database into tables that match the appropriate RDF class structure, and then convert the data in those tables into the appropriate format with the relational DB → RDF converter.
 - 2. The above process might work in both directions, i.e. it might work for both BioWarehouse → BioPAX conversion and for BioPAX → BioWarehouse conversion.
 - iii. Mike suggested Jeremy contact Peter about his ideas
- 3. Best practice recommendation on BioPAX data access
 - a. Mike pointed out that we currently provide no recommendations to data providers on how they should make their data available, and as a result it looks like the first three DBs to make their data available in BioPAX format will do so in different ways:
 - i. BioCyc provides BioPAX data via flat-files that are available via FTP. They create one BioPAX file per organism, each file is named 'biopax.owl'
 - 1. Joanne suggested naming each file by organism, e.g. 'hpycyc.biopax'
 - ii. KEGG – the organization that appears most likely at this point is 1 BioPAX file per pathway map per organism (e.g. glycolysis in human would be one file, glycolysis in E. coli would be another).
 - iii. WIT – up to this point they have been organizing their data by class, sending Jeremy (for him to review) 1 file per major class type (e.g. 1 file containing all small molecules in their DB, 1 file containing all reactions, etc.)
 - b. Group consensus: develop recommendations to data providers on how they should make BioPAX data available, or at least minimum recommended standard for data access and availability
 - i. E.g. "Provide data as flat-files available for download, organized by organism"
 - ii. Jeremy suggested recommending users label their BioPAX data with LSID's corresponding to the original data
 - iii. Mike and Gary will set up a wiki page with a draft recommendation
- 4. Wiki
 - a. Recent content updates (new pages):
 - i. Authorship guidelines
 - ii. Sample use cases
 - iii. Development process and release checklist
 - 1. Jonathan Rees has contributed a number of updates to these
 - b. Mike will send out an announcement about the wiki in the near future and link to it from the biopax.org homepage
- 5. Grants

- a. DOE extension and renewal
 - i. Chris has a call scheduled to John Houghton on Friday to discuss progress and next steps
 - ii. Addendum: This call did not occur, it will be rescheduled.
- b. NIGMS conference grant
 - i. Mike will follow up with them in the near future to find out award status
- 6. Next BioPAX face-to-face meeting: June 25 in Detroit (at ISMB)
 - a. Location: Richards rooms 1 & 2 in Marriott Renaissance Center
 - b. Time: 10 AM – 5 PM
 - c. Agenda:
 - i. Discuss representation of biological states – States subgroup developing a proposal (see states yahoo group for details: <http://groups.yahoo.com/group/biopax-states/>)
 - ii. Discuss representation of gene regulation – Mike and Gary will develop proposal in coming weeks
 - iii. (Maybe, if time) Discuss representation of generic entities – possible representation being developed by states subgroup (potential simple solution came to light while discussing states)
 - iv. Discuss goals, requirements, and development process for Level 3
 - v. Ratify Level 2
 - vi. Wiki Q&A session
 - 1. Discuss issues raised on wiki (e.g. questions posted by users)
 - d. Other BioPAX events at ISMB:
 - i. Poster at Bio-Ontologies SIG
 - ii. Talk at BioPathways SIG
 - iii. Poster at main conference poster session
 - iv. BOF meeting – 1 hour open meeting during lunch
 - 1. Group consensus: Sunday would be best day
 - v. Panel discussion at end of BPC meeting
 - 1. Not confirmed, schedule for BPC meeting still under development
- 7. Other future meetings
 - a. BioPAX-Boston meeting: Friday, June 3, 2005
 - i. Smaller group than previous BioPAX-Boston meeting (so it will be more focused on BioPAX issues rather than general pathway issues)
 - ii. Meeting focus: meeting data conversion objectives by ISMB deadline
 - iii. Additional possible topic: SBML and BioPAX integration
 - 1. Previous work focused on using BioPAX to annotate SBML
 - a. Using SBML's metadata tag
 - 2. New proposal: convert SBML to RDF, mix and match BioPAX and SBML RDF
 - a. Jeremy discussed the possibility with Ben Bornstein at the SBML hackathon in Tokyo

