Table of Contents

Keynote Presentations

Enabling the Molecular Medicine Revolution Through Network-Centric Biomedicine .................................................. 1
Kenneth H. Buetow

Phyl-O’Data (POD) from Tree of Life: Integration Challenges from Yellow Slimy Things to Black Crunchy Stuff .................. 3
Junhyong Kim

New Architectures and Experience on Using Systems

Automatically Constructing a Directory of Molecular Biology Databases ................................................................. 6
Luciano Barbosa, Sumit Tandon, and Juliana Freire

The Allen Brain Atlas: Delivering Neuroscience to the Web on a Genome Wide Scale .................................................. 17
Chinh Dang, Andrew Sodt, Chris Lau, Brian Youngstrom, Lydia Ng, Leonard Kuan, Sayan Pathak, Allan Jones, and Mike Hawrylycz

Toward an Integrated RNA Motif Database ................................................. 27
Jason T.L. Wang, Dongrong Wen, Bruce A. Shapiro, Katherine G. Herbert, Jing Li, and Kaushik Ghosh

B-Fabric: A Data and Application Integration Framework for Life Sciences Research ....................................................... 37
Can Türker, Etzard Stolte, Dieter Joho, and Ralph Schlapbach

SWAMI: Integrating Biological Databases and Analysis Tools Within User Friendly Environment ........................................... 48
Rami Rifaieh, Roger Unwin, Jeremy Carver, and Mark A. Miller

myGrid and UTOPIA: An Integrated Approach to Enacting and Visualising in Silico Experiments in the Life Sciences ................. 59
Steve Pettifer, Katy Wolstencroft, Pinar Alper, Teresa Attwood, Alain Coletta, Carole Goble, Peter Li, Philip McDermott, James Marsh, Tom Oinn, James Sinnott, and David Thorne

Managing and Designing Scientific Workflows

A High-Throughput Bioinformatics Platform for Mass Spectrometry-Based Proteomics .................................................. 71
Thodoros Topaloglou, Moyez Dharsee, Rob M. Ewing, and Yury Bukhman
Bioinformatics Service Reconciliation by Heterogeneous Schema Transformation ......................................................... 89
   Lucas Zamboulis, Nigel Martin, and Alexandra Poulovassilis

A Formal Model of Dataflow Repositories .............................................. 105
   Jan Hidders, Natalia Kwasnikowska, Jacek Sroka,
   Jerzy Tyszkiewicz, and Jan Van den Bussche

Project Histories: Managing Data Provenance Across Collection-Oriented Scientific Workflow Runs .................................................. 122
   Shawn Bowers, Timothy McPhillips, Martin Wu, and
   Bertram Ludäscher

Mapping and Matching Techniques

Fast Approximate Duplicate Detection for 2D-NMR Spectra ..................... 139
   Björn Egert, Steffen Neumann, and Alexander Hinneburg

Ontology-Supported Machine Learning and Decision Support in Biomedicine ................................................................. 156
   Alexey Tsymbal, Sonja Zühlker, and Martin Huber

Instance-Based Matching of Large Life Science Ontologies ...................... 172
   Toralf Kirsten, Andreas Thor, and Erhard Rahm

Modeling of Life Science Data

Data Integration and Pattern-Finding in Biological Sequence with TESS’s Annotation Grammar and Extraction Language (AnGEL) ................................. 188

Inferring Gene Regulatory Networks from Multiple Data Sources Via a Dynamic Bayesian Network with Structural EM ...................................... 204
   Yu Zhang, Zhidong Deng, Hongshan Jiang, and Peifa Jia

Accelerating Disease Gene Identification Through Integrated SNP Data Analysis ..................................................................................... 215
   Paolo Missier, Suzanne Embury, Conny Hedeler, Mark Greenwood,
   Joanne Pennock, and Andy Brass

Annotation in Data Integration

What’s New? What’s Certain? – Scoring Search Results in the Presence of Overlapping Data Sources ............................................................ 231
   Philipp Hussels, Silke Trißl, and Ulf Leser

Using Annotations from Controlled Vocabularies to Find Meaningful Associations ...................................................................................... 247
   Woei-Jyh Lee, Louiza Raschid, Padmini Srinivasan, Nigam Shah,
   Daniel Rubin, and Natasha Noy