

Agenda
5th Meeting on U.S. Government Chemical Databases and Open Chemistry

Thursday, August 25, 2011

MORNING

Chairperson: Ann Richard

9:00-9:15 AM

Registration, Check-In

9:15-9:30 AM

Introduction: Marc Nicklaus

9:30-11:15 AM

Morning Session I – FDA Databases & Projects

9:30-9:50

Kirk Arvidson, FDA/CFSAN, “Development of an Institutional Knowledge-base at FDA’s Center for Food Safety and Applied Nutrition”

9:50-9:55

Discussion

9:55-10:15

Chihae Yang, Altamira LLC, “Ontology Development for Nanomaterials Safety”

10:15-10:20

Discussion

10:20-10:40

Huixiao Hong, FDA/NCTR, “EDKB: Endocrine Disruptors Knowledge Base at the FDA”

10:40-10:45

Discussion

10:45-11:05

Frank Switzer, FDA/OC, “The FDA/USP Substance Registration System as an Informational Bridge Between Chemistry and Biology”

11:05-11:15

Discussion

11:15-11:30 AM

Morning Break

11:30-12:45 PM

Morning Session II – Toxicology

11:30-11:50

Eugene Demchuk, CDC, “Chemical Toxicity Databases at NCEH/ATSDR”

11:50-11:55

Discussion

11:55-12:15

Ann Richard, EPA/NCCT, “Tox21 and ToxCast Chemical Landscapes: Laying the Foundation for 21st Century Toxicology”

12:15-12:20

Discussion

12:20-12:40

Nina Jeliaskova, Ideaconult Ltd., “Seamless and uniform access to chemical data and tools: Experience gained in developing the OpenTox framework”

12:40-12:45

Discussion

12:45-1:45 PM

Lunch Break & Poster Session

Thursday, August 25, 2011

AFTERNOON

Chairperson: Michael Gilson

1:45-2:45 PM

Afternoon Session I – Identifiers

1:45-2:05

Steve Heller, NIST, “The Status of the InChI Project”

2:05-2:15

Discussion

2:15-2:35

Jonathan Goodman, IUPAC & RSC, “InChI and Reactions”

2:35-2:45

Discussion

2:45-3:00 PM

Afternoon Break

3:00-4:45 PM

Afternoon Session II – NIH Resources and How to Access Them

3:00-3:20

Evan Bolton, NCBI/NLM, “PubChem: A significant resource for scientists”

3:20-3:25

Discussion

3:25-3:45

Wolf-D. Ihlenfeldt, Xemistry GmbH, “Accessing U.S. Government Databases with the CACTVS Toolkit”

3:45-3:50

Discussion

3:50-4:10

Markus Sitzmann, NCI, “NCI/CADD Chemical Identifier Resolver: Indexing and Analysis of Available Chemistry Space”

4:10-4:15

Discussion

4:15-4:35

Marcus Hanwell, Kitware, Inc., “Chemical Databases and Open Chemistry on the Desktop”

4:35-4:45

Discussion

4:45-6:00 PM

Panel Discussion

Topic: “How do we get (authors to put, and publishers to accept if not require) chemistry information in a computer-readable format into primary chemical literature?”

Jasmine Young

Rutgers; RCSB PDB

Michael Gilson

UCSF; Binding DB database

Steve Heller

NIST; InChI

Dušanka Janežič

NIC, Ljubljana; JCIM, Associate Editor

Jay Schneekloth

CBL, NCI; Synthetic Chemistry & Screening

Markus Sitzmann

CADD Group, NCI; Chem. Identifier Resolver

Antony Williams

RSC; ChemSpider

Friday, August 26, 2011

MORNING

Chairperson: Heather Carlson

9:00-9:15 AM

Registration, Check-In

9:15-11:00 AM

Morning Session I – Ligands and Binding Data

9:15-9:35

Jasmine Young, RCSB, “Management and Distribution of Chemical Data in the PDB”

9:35-9:40

Discussion

9:40-10:00

James Dunbar, UMich, “CSAR and Binding MOAD: Two different databases, two different aims, one common goal provide the best protein-ligand data”

Discussion

10:00-10:05

10:05-10:25

Renxiao Wang, Shanghai Inst. Org. Chem., “The PDBbind Database: A Comprehensive Collection of the Binding Data and Structures of the Complexes in the Protein Data Bank”

Discussion

10:25-10:30

10:30-10:50

Michael Gilson, UCSD, “BindingDB: A Protein-Ligand Database for Drug Discovery”

Discussion

10:50-11:00

11:00-11:15 AM

Morning Break

11:15-12:30 AM

Morning Session II – Drug Design; Validation of Chemistry Data

11:15-11:35

John Overington, EMBL-EBI, “ChEMBL - An Open Data resource for drug discovery”

11:35-11:40

Discussion

11:40-12:00

Antony Williams, RSC, “ChemSpider: A Crowdsourcing Environment for Hosting and Validating Chemistry Resources”

Discussion

12:00-12:05

12:05-12:25

Talapady N Bhat, NIST, “Managing 2D and 3D Ligand Fragments from Disparate Large Resources”

Discussion

12:25-12:30

12:30-1:30 PM

Lunch Break & Poster Session

Friday, August 26, 2011

AFTERNOON

Chairperson: Markus Sitzmann

1:30-4:00 PM

Afternoon Session I – Drug Design; Validation of Chemistry Data

1:30-1:50

John Irwin, UCSF, “ZINC – A database of commercially available compounds for virtual screening”

1:50-1:55

Discussion

1:55-2:15

Scott Hutton, ChemNavigator/Sigma-Aldrich, “Aldrich Market Select Chemistry Portal, a new publicly accessible resource under development at Sigma-Aldrich”

2:15-2:20

Discussion

2:20-2:50

Noel O’Boyle, Open Babel Development Team, “Improving the quality of chemical databases with community-developed tools (and vice versa)”

2:50-2:55

Discussion

2:55-3:15

Andrew Lang, Oral Roberts Univ., and Jean-Claude Bradley, Drexel Univ., “The collection, curation and modeling of open melting point measurements”

3:15-3:20

Discussion

3:20-3:50

Martin Walker, SUNY, “Validation of Wikipedia Chemical Data”

3:50-4:00

Discussion

4:00-4:15 PM

Afternoon Break

4:15-5:55 PM

Afternoon Session II – Open Tools; OSR and Patents

4:15-4:35

Gregory Landrum, NIBR, “The RDKit and PostgreSQL: An open-source database system for chemistry”

4:35-4:40

Discussion

4:40-5:00

Koji Nakagawa, Kyushu Univ., “Digitizing Chemical Structure Images of Japanese Published Patent Applications into a Searchable Format”

5:00-5:05

Discussion

5:05-5:25

Marc Zimmermann, Univ. Bonn, “Evaluation of different benchmark sets and evaluation methods for automatic extraction of chemical entities from text and image”

5:25-5:30

Discussion

5:30-5:50

Stephen Boyer, IBM, “Computer Curation of WW Intellectual Property (patents) and the Scientific Literature”

5:50-5:55

Discussion

5:55-6:00 PM

Concluding Remarks

POSTER SESSION

- 1. XML-based description and uniqueness check for complex substances in the FDA Substance Registration System - Or: How to assure a UNII is actually UNIIque**
Yulia Borodina, Frank Switzer, Tyler Peryea, and Lawrence Callahan
OCPP, OCS, OC, FDA
- 2. ToxML: A Format for Toxicology Data and Analysis Results -**
R. Daniel Benz¹, Kevin P. Cross² and David A. Bower²
¹FDA/CDER/OTR/DDSR and ²Leadscope, Inc
- 3. Chemical Evaluation and Risk Estimation System (CERES) Data Model Overview**
Dimitar Hristozov, Annette McCarthy, Kirk Arvidson, and Chihae Yang
Office of Food Additive Safety, Center for Food Safety and Applied Nutrition,
U.S. Food and Drug Administration, College Park, MD
- 4. FDI: Flexible Database Integrator for integrating multiple databases**
Uma Mudunuri, Anney Che and Robert Stephens
Advanced Biomedical Computing Center, SAIC-Frederick, NCI-Frederick, MD
- 5. A Suite of Web Resources for Computational RNA Analysis, Prediction and Design**
Eckart Bindewald¹, Wojciech Kasprzak¹, Bruce A. Shapiro²
¹Basic Science Program, SAIC-Frederick, Inc., NCI-Frederick, Frederick,
Maryland, ²Center for Cancer Research Nanobiology Program, NCI-Frederick, MD
- 6. NIST Computational Chemistry Comparison and Benchmark Database**
Russell Johnson
NIST, Gaithersburg, MD
- 7. Automating the process of 3D structure generation for 2D registration into a chemical cartridge database**
Mitchell Miller¹, George (Mike) Hazard², Chuchu Lan¹
¹Scientific Thinking, LLC, Fairfax, VT. ²National Library of Medicine,
Specialized Information Services, Bethesda, MD
- 8. The NIAID ChemDB HIV/AIDS Database**
Margaret Rush¹, Danna Huffman², Glen Noble¹, Matthew Whiting¹, Mohamed Nasr³
¹Gryphon Scientific LLC, Takoma Park, MD; ²SRA International, Inc., Fairfax, VA;
³NIH, NIAID, Bethesda, MD
- 9. Development of Virtual Screening Models Based on NIH Roadmap Assay Data - Experimental Validation**
Iwona E Weidlich^{1,2}, Igor Filippov³, and Marc C Nicklaus¹
¹Chemical Biology Laboratory, Center for Cancer Research, NCI, NIH, DHHS,
Frederick, MD; ²University of Maryland, Baltimore County, Chemistry and
Biochemistry Department, 1000 Hilltop Circle, Baltimore, MD; ³Chemical Biology
Laboratory, SAIC-Frederick, Inc., NCI-Frederick, MD

10. Chemoinformatics infrastructure for drug discovery

Kamal Kumar, Li Cheng, Jason Smith, Sid Chaudhury, Narender Singh, Mohamed Diwan AbdulHameed, Gregory Tawa, and Anders Wallqvist
Biotechnology HPC Software Applications Institute, Telemedicine and Advanced Technology Research Center, U.S. Army Medical Research and Materiel Command, Fort Detrick, MD

11. A Physicochemical Descriptor-based Scoring Scheme for Effective and Rapid Filtering of Kinase-like Chemical Space Extracted from the ChEMBL Database

Narender Singh, Sidhartha Chaudhury, Mohamed Diwan M. AbdulHameed, Anders Wallqvist, and Gregory Tawa
DoD Biotechnology High Performance Computing Software Applications Institute, Telemedicine and Advanced Technology Research Center, U.S. Army Medical Research and Materiel Command, Fort Detrick, MD

12. BindingDB: A Protein-Ligand Database for Drug Discovery

George Nicola, Tiqing Liu, Linda Hwang and Michael K. Gilson
University of California San Diego

13. Databases and Tools Developed by LIPID MAPS Consortium

Manish Sud^a, Eoin Fahy^a, Dawn Cotter^a, Robert Byrnes^a, Shakti Gupta^a, Mano Maurya^a, H. Alex Brown^c, Christopher K. Glass^d, Alfred H. Merrill, Jr.^e, Robert C. Murphy^f, Christian R.H. Raetz^g, David W. Russell^h, Edward A. Dennisⁱ, Shankar Subramaniam^{a,b}
LIPID MAPS Bioinformatics Core^a, SDSC/UCSD, 9500 Gilman Dr, La Jolla, CA 92093; Department of Bioengineering^b, UCSD, 9500, Gilman Dr, La Jolla, CA, 92093; Department of Pharmacology^c, Vanderbilt University Medical Center, Nashville, TN 37232; Department of Cellular and Molecular Medicine^d, UCSD, 9500 Gilman Dr, La Jolla, CA 92093; School of Biology^e, Georgia Institute of Technology, Atlanta, GA 30332; University of Colorado Health Sciences Center^f, Aurora, CO 80045; Department of Biochemistry^g, Duke University Medical Center, Durham, NC 27710; Department of Molecular Genetics^h, University of Texas Southwestern Medical Center, Dallas, Texas 75390; Department of Chemistry and Biochemistry, and Department of Pharmacologyⁱ, UCSD, La Jolla, CA 92093.

14. Chemical eCommerce

Klaus Gubernator
eMolecules, Inc.

15. Optical Structure Recognition Application: development timeline and the status quo

Igor Filippov¹, and Marc C Nicklaus²
¹Chemical Biology Laboratory, SAIC-Frederick, Inc., NCI-Frederick, MD; ²Chemical Biology Laboratory, Center for Cancer Research, NCI, NIH, DHHS, Frederick, MD

16. Image-Based Automated Annotation of Chemical Database: Development and Application

Jungkap Park¹, Gus Rosania², and Kazuhiro Saitou¹.
¹Department of Mechanical Engineering, University of Michigan, Ann Arbor, MI;
²Department of Pharmaceutical Sciences, University of Michigan, Ann Arbor, MI

17. On the Development of OCR Software for Chemical Structure Images in Japanese Published Patent Applications

Akio Fujiyoshi¹, Koji Nakagawa², and Masakazu Suzuki².

¹Faculty of Engineering, Ibaraki University, Hitachi, Ibaraki, 316-8511 Japan;

²Faculty of Mathematics, Kyushu University, Nishi-ku, Fukuoka, 819-0395 Japan