<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Message From The Chair</td>
<td>3</td>
</tr>
<tr>
<td>Letter From The Editor</td>
<td>5</td>
</tr>
<tr>
<td>ACS CINF Social Networking and Business Events at the Fall 2013 ACS Meeting</td>
<td>6</td>
</tr>
<tr>
<td>Awards and Scholarships</td>
<td>10</td>
</tr>
<tr>
<td>Committee Reports</td>
<td>17</td>
</tr>
<tr>
<td>Interview with Maureen Rouhi, Editor-in-Chief of Chemical &amp; Engineering News by Svetla Baykoucheva</td>
<td>20</td>
</tr>
<tr>
<td>Book Reviews by Bob Buntrock</td>
<td>25</td>
</tr>
<tr>
<td>Literature Digest by Song Yu</td>
<td>27</td>
</tr>
<tr>
<td>CINF Reader Submitted Chemical Information Topic of Interest — The Growth of Semantic Technologies by Vijay Bhatia</td>
<td>29</td>
</tr>
<tr>
<td>Notes From Our Sponsors</td>
<td>33</td>
</tr>
<tr>
<td>Fall 2013 ACS CINF and Cosponsored Symposia Technical Program</td>
<td>39</td>
</tr>
<tr>
<td>Fall 2013 ACS CINF Abstracts</td>
<td>65</td>
</tr>
<tr>
<td>2013 CINF Officers and Functionaries</td>
<td>127</td>
</tr>
<tr>
<td>Future ACS Meetings</td>
<td>130</td>
</tr>
</tbody>
</table>
Message From the Chair

Indianapolis, ahoy! So I won’t be sailing into Indianapolis for the next ACS meeting, I’ll be flying, but I am excited to show up for a conference that is sure to stimulate, educate and yes, tire me out. Every ACS meeting is stacked end-to-end with talks that I want to get to, events that I want to experience, new people I want to network with and, of course, fellow CINF’ers with whom I want to catch up and see what’s happening in our ever expanding world of chemical information. Since this meeting is dedicated to “Chemistry in Motion,” we intend to ensure that you want to keep moving between symposia and social events! Indy offers us a rich program with many of you chairing, speaking in, and/or attending our CINF sessions. The program is already online and you can check out what’s coming here:

http://abstracts.acs.org/chem/246nm/program/divisionindex.php?act=session&val=191016&prog=191016

The sessions run from Sunday through Thursday and, having been involved in organizing a number of them myself, I can tell you that there has been a lot of effort put into allocating the talks to the best sessions and making sure that the days are full. We are even involved in 9 additional sessions co-sponsored by other divisions (CHAS and COMP). At the upcoming meeting you will have the chance to hear about the latest and greatest advances on Tablet Computers, sit in on a day dedicated to Dick Cramer, the winner of our 37th Herman Skolnik award, and discuss budding research with graduate students in cheminformatics and information science. These are just some of what’s on offer and we hope that our CINF programming committee has done you proud with what you will experience. We will of course have a number of social events and will do our utmost to ensure that the food and drink is both on time and up to expectations!

Over the past few months we have tried to extend ourselves a little more as a Division into the world of social networking. Hopefully some of you have noticed. Our greatest success has likely been the delivery of a number of webinars (ably hosted by our own David Martinsen!), posted online here:

http://www.youtube.com/user/cinfacs?feature=watch

In my opinion we are still struggling to reach outside our Division to further encourage membership but, more importantly, to help the chemistry community in general get more of a grip on what is possible today with chemical information. We are challenged by a growing “Google it” mentality (which is incredibly useful by the way) relative to the advantages that a good grounding in chemical information skills can bring to students and career scientists alike. Part of my personal mission is to spread the message regarding social networking skills for scientists not only to help them in the representation of their own profiles as scientists, but also to help them network into our world. My request to you as CINF members is pass on to your colleagues outside the Division those nuggets that might show up on the CHMINF list. Invite them to our webinars. Point them to our website. Invite them along to our symposia and our social events and, ultimately, give them a chance to embrace what we have to offer them.
So, that's Indianapolis on the horizon. Even before that, the call for papers for Dallas Spring 2014 goes live. We are presently rounding out the program and look forward to unveiling it. I’m looking forward to seeing you all in Indy in September. If you can't make it we will do our best to summarize the events in our next *Chemical Information Bulletin*! Hope you all have a great summer!

Dr. Antony J. Williams  
Chair, ACS Division of Chemical Information  
Tony27587@gmail.com
Letter From The Editor

Thanks for reading the ACS Chemical Information Bulletin (CIB). I am a new member of the CINF community and it has been a fantastic experience contributing a small part to the CINF Division as editor of the Fall 2013 CIB. The previous ACS meeting in New Orleans (245th) was my first experience with the CINF Division. Wow, what a fantastic, dedicated, and welcoming group. If you are not a CINF member already, I encourage you to check us out!

You may notice the new look of the CIB. We were able to take the ACS Adobe InDesign stationary templates and modify them for the CIB. The new design now better reflects the recommended ACS graphic standards. More importantly, I hope the new design brings a modern, consistent, and more simplified digital look to the PDF version of the CIB. We hope you enjoy the new face of the CIB, and welcome any feedback. I’m not sure how or even why, but typography has become somewhat of a new hobby for me, so laying out the CIB in Adobe InDesign was a lot of fun (though I have much to learn, so for all the real typographers and graphic designers out there, go easy on me)! We also hope to bring a complete epub version of the CIB alongside the PDF and HTML versions in the near future.

In this pre-conference issue, you will find two book reviews by Bob Buntrock, an interview with Maureen Rouhi (Editor-in-Chief of Chemical & Engineering News) by Svetla Baykoucheva, and recommended literature reading by Song Yu. We have also included news and report updates from a subset of our committees, notes from our generous sponsors, and the technical program for the upcoming 246th ACS Meeting in Indianapolis, IN (Sept. 8–12). In addition, one of our readers (and future CINF member) submitted a short summary of his recent interest with linked data and the semantic web.

I would like to thank everyone who submitted content and assisted with the production of the CIB. And many thanks to Svetla Baykoucheva for introducing me to David Martinsen and recommending the Editor position. Lastly, I will be editing the fall and spring pre-conference Issues of the CIB, while Svetlana Korolev will continue as the post-conference Winter and Summer CIB Editor.

Vincent F. Scalfani, Editor
vfscalfani@ua.edu
ACS CINF Social Networking Events at the Fall 2013 ACS Meeting

Please Join Us at these Division of Chemical Information Events!

The ACS Division of Chemical Information is pleased to host the following social networking events at the Fall 2013 ACS National Meeting in Indianapolis, IN.

**CSA Trustees**
12:00-2:00 pm, Sunday, September 8th – JW Room 106, JW Marriott.

**Sunday Welcoming Reception & CINF Scholarships for Scientific Excellence Posters**
6:30-8:30 pm, Sunday, September 8th – Wabash Ballroom Section 1/2, Indiana Convention Center. Reception co-sponsored by Bio-Rad Laboratories, InfoChem, Optibrium, PerkinElmer, Thieme Chemistry, and ACS Symposium Series.

Scholarships for Scientific Excellence sponsored exclusively by RSC Publishing.

**Social Networking Workshop**
4:00-5:30 pm, Monday, September 9th – ICC 141.

**Harry’s Party**
5:30-8:00 pm Monday, September 9th – Governor’s Suite, Westin Indianapolis Hotel.
Sponsored exclusively by ACS Publications. *Skywalk from Indiana Convention Center*

**Tuesday Luncheon** (Ticketed Event – Contact Division Chair, Tony Williams)
12:00-1:30 pm Tuesday, September 10th – Wabash Ballroom Section 2, Indiana Convention Center.
Sponsored exclusively by the Royal Society of Chemistry.
*Speaker: Katy Börner - Multi-Scale Maps of Scholarly Activity (abstract on p. 8)*

**Herman Skolnik Award Reception honoring Dr. Richard (Dick) Cramer**
6:30-8:30pm Tuesday, September 10th – JW Grand Ballroom 4, JW Marriott Indianapolis.
ACS CINF Business Events at the Fall 2013 ACS Meeting

Saturday, September 7.

1. Awards Committee, 1-3 pm, ICC 110
2. Education Committee, 1-3 pm, ICC 111
3. Program Committee, 1-3 pm, ICC 112
4. Executive Committee, 3-6 pm, ICC 111
Tuesday Luncheon Talk – Multi-Scale Maps of Scholarly Activity – KATY BÖRNER

Cartographic maps of physical places have guided mankind’s explorations for centuries. They enabled the discovery of new worlds while also marking territories inhabited by unknown monsters. Domain maps of abstract topic spaces aim to provide guidance for navigating, understanding, and communicating the dynamic and changing structure of science and technology. The maps are generated through scientific analysis of large-scale scholarly data sets in an effort to connect and make sense of the bits and pieces of knowledge they contain.

The Mapping Science exhibit (http://scimaps.org) aims to introduce science mapping techniques to the general public and to experts across diverse disciplines for educational, scientific, and practical purposes. It is meant to inspire crossdisciplinary discussion on how to best track and communicate scholarly activity and scientific progress on a global scale. Science maps serve as visual interfaces to immense amounts of data—depicting perhaps millions of objects that we can perceive rapidly to effectively discern apparent outliers, clusters, and trends.

This talk will present a visual feast of science maps, data mining and visualization tools such as plug-and-play macroscopes (http://cishell.org), and the Information Visualization MOOC (http://ivmooc.cns.iu.edu) freely available to anyone interested to derive insights from large-scale datasets.

References:
KATY BÖRNER is the Victor H. Yngve Professor of Information Science at the Department of Information and Library Science at the School of Informatics and Computing, Adjunct Professor at the Department of Statistics in the College of Arts and Sciences, Core Faculty of Cognitive Science, Research Affiliate of the Biocomplexity Institute, Fellow of the Center for Research on Learning and Technology, Member of the Advanced Visualization Laboratory, and Founding Director of the Cyberinfrastructure for Network Science Center (http://cns.iu.edu) at Indiana University. She holds a Visiting Professor appointment at the Royal Netherlands Academy of Arts and Sciences (KNAW), The Netherlands. Since 2004, she has been a curator of the Places & Spaces: Mapping Science exhibit (http://scimaps.org).

Her research focuses on the development of data analysis and visualization techniques for information access, understanding, and management. She is particularly interested in the study of the structure and evolution of scientific disciplines; the analysis and visualization of online activity; and the development of cyberinfrastructures for large scale scientific collaboration and computation. Börner is the co-editor of the books ‘Visual Interfaces to Digital Libraries’ (2003) and ‘Modeling Science Dynamics’ (2012), ‘VIVO: A Semantic Approach to Scholarly Networking and Discovery’ (2012), and of a special PNAS issue on ‘Mapping Knowledge Domains’ (2004). Her book ‘Atlas of Science: Visualizing What We Know’ by MIT Press was published in 2010 and won the ASIS&T best Information Science book award in 2011. She holds a MS in Electrical Engineering from the University of Technology in Leipzig, 1991 and a Ph.D. in Computer Science from the University of Kaiserslautern, 1997.
Bonnie Lawlor to be Inducted as 2013 ACS Fellow

Bonnie Lawlor, CINF Councilor, will be inducted as an ACS Fellow, class of 2013, during the fall National Meeting of the American Chemical Society (ACS) in Indianapolis. Bonnie began her career at the Institute for Scientific Information (now part of Thomson Reuters) as a chemical indexer/editor for Index Chemicus (IC). Her leadership ability became very apparent and she quickly rose to manager of the department. She eventually became an ISI Director and was responsible for the overall production of the chemical products including the IC Registry System (ICRS). In the mid-1980’s she was promoted to Vice President. In this position the staff and all production phases of Current Contents (CC) were added to her responsibilities. Bonnie demonstrated a take-charge attitude and was eventually given the responsibility of content management for the all of ISI’s products including the Science Citation Index (SCI), Social Science Citation Index (SSCI), and Arts and Humanities Citation Index (AHCI). Through various promotions she became Executive Vice President of database publishing. In addition to overseeing production of the many ISI products she was given the added responsibility of global sales and marketing.

When new technologies were developed for delivering products Bonnie was at the forefront for pursuing how these technologies could best be used for ISI products. She was involved in making the relevant products available first on diskette and then on CD ROM which became the forerunner for the Web of Science (WOS) which is currently used by millions of students and researchers in colleges and universities throughout the world. The WOS database is also widely used in evaluative and scientometrics studies.

Bonnie is currently Executive Director of the National Federation of Advanced Information Services (NFAIS), a membership organization of more than 55 of the world’s leading producers of databases and related information services, information technology, and library services in the sciences, engineering, social sciences, business, and the arts and humanities.

Bonnie became a member of the American Chemical Society in 1972 and almost immediately became an active member. As a member of the Division of Chemical Information she has had many roles. She started out as Assistant Editor of the CIB and then became Editor. She was Secretary (1984-1987), Chair-Elect (1988), Chair (1989), Past-Chair (1990) and is currently Councilor, a position she has held for more than 20 years. Bonnie has been a member of many committees and served as chair for several of them. She is currently the CINF Archivist and is on the CINF Communications and Publications committee. As a Councilor she has held many elected positions at the national level. Some of these are Committee on Committees (2012-present), Council Policy Committee (voting 2006-2011, nonvoting 1997-1999), Committee on Nominations & Elections (2000-2005, vice-chair 2003, secretary 2001), Committee on Divisional Activities (1994-1999, chair 1997-1999), and Committee on Copyrights (1990-1998, Chair 1993-1995, committee associate 1989).
In 2010 Bonnie was interviewed by the then editor of the *CIB* Svetla Baykoucheva. That interview can be found at [http://bulletin.acscinf.org/node/188](http://bulletin.acscinf.org/node/188).

Above is only a snapshot of Bonnie’s career accomplishments and her many contributions to ACS. This is enough to give you a sense of why CINF’s Bonnie Lawlor was selected as one of the 2013 ACS Fellows.

Margaret Matthews
Thomson Reuters

**Peter Rusch to be Inducted as 2013 ACS Fellow**

Peter Rusch, a past CINF chair, and longtime contributor to the CINF division will also be inducted as a 2013 ACS Fellow. See a brief outline of his contributions to science and the ACS Community here:

Applications Invited for CSA Trust Grants for 2014

The Chemical Structure Association (CSA) Trust is an internationally recognized organization established to promote the critical importance of chemical information to advances in chemical research. In support of its charter, the Trust has created a unique Grant Program and is currently inviting the submission of grant applications for 2014.

Purpose of the Grants

The Grant Program has been created to provide funding for the career development of young researchers who have demonstrated excellence in their education, research or development activities that are related to the systems and methods used to store, process and retrieve information about chemical structures, reactions and compounds. One or more Grants will be awarded annually up to a total combined maximum of ten thousand U.S. dollars ($10,000). Grants are awarded for specific purposes, and within one year each grantee is required to submit a brief written report detailing how the grant funds were allocated. Grantees are also requested to recognize the support of the Trust in any paper or presentation that is given as a result of that support.

Who is Eligible?

Applicant(s), age 35 or younger, who have demonstrated excellence in their chemical information related research and who are developing careers that have the potential to have a positive impact on the utility of chemical information relevant to chemical structures, reactions and compounds, are invited to submit applications. While the primary focus of the Grant Program is the career development of young researchers, additional bursaries may be made available at the discretion of the Trust. All requests must follow the application procedures noted below and will be weighed against the same criteria.

Which Activities are Eligible?

Grants may be awarded to acquire the experience and education necessary to support research activities; e.g. for travel to collaborate with research groups, to attend a conference relevant to one’s area of research, to gain access to special computational facilities, or to acquire unique research techniques in support of one’s research.

Application Requirements

Applications must include the following documentation:

1. A letter that details the work upon which the Grant application is to be evaluated as well as details on research recently completed by the applicant;
2. The amount of Grant funds being requested and the details regarding the purpose for which the Grant will be used (e.g. cost of equipment, travel expenses if the request is for financial support of meeting attendance, etc.). The relevance of the above-stated purpose to the Trust's objectives and the clarity of this statement are essential in the evaluation of the application;

3. A brief biographical sketch, including a statement of academic qualifications;

4. Two reference letters in support of the application.

Additional materials may be supplied at the discretion of the applicant only if relevant to the application and if such materials provide information not already included in items 1-4. Three copies of the complete application document must be supplied for distribution to the Grants Committee.

**Deadline for Applications**

Applications must be received no later than March 13, 2014. Successful applicants will be notified no later than May 2, 2014.

**Address for Submission of Applications**

Three copies of the application documentation should be forwarded to: Bonnie Lawlor, CSA Trust Grant Committee Chair, 276 Upper Gulph Road, Radnor, PA 19087, USA. If you wish to enter your application by e-mail, please contact Bonnie Lawlor at chescot@aol.com prior to submission so that she can contact you if the e-mail does not arrive.

**Chemical Structure Association Trust: Recent Grant Awardees**

**2013 – Dr. Johannes Hachmann***

Department of Chemistry and Chemical Biology at Harvard University, Cambridge, MA. He was awarded the Grant for travel to speak on “Structure-property relationships of molecular precursors to organic electronics at a workshop sponsored by the Centre Européen de Calcul Atomique et Moléculaire (CECAM) that will take place October 22 – 25, 2013 in Lausanne, Switzerland.

**2013 – Dr. Robert S. Paton***

University of Oxford, UK. He was awarded the Grant to speak at the Sixth Asian Pacific Conference of Theoretical and Computational Chemistry in Korea on July 11, 2013. Receiving the invitation for this meeting has provided Dr. Paton with an opportunity to further his career as a Principal Investigator.

**2013 – Dr. Aaron Thornton***

Material Science and Engineering at CSIRO in Victoria, Australia. He was awarded the Grant to attend the 2014 International Conference on Molecular and Materials Informatics at Iowa State University with the objective of expanding his knowledge of web semantics, chemical mark-up language, resource description frameworks and other online sharing
tools. He will also visit Dr. Maciej Haranczyk, a prior CSA Trust Grant recipient, who is one of the world leaders in virtual screening.

2012 – Tu C. Le

CSIRO Division of Materials Science & Engineering, Clayton, VIV, Australia. Tu C. was awarded the Grant for travel to attend a Cheminformatics course at Sheffield University and to visit the Membrane Biophysics group of the Department of Chemistry at imperial College London.

2011 – J. B. Brown

Kyoto University, Kyoto, Japan. J.B. was awarded the Grant for travel to work with Professor Ernst Walter-Knappat the Freie University of Berlin and Professor Jean-Phillipe Vert of the Paris MinesTech to continue his work on the development of atomic partial charge kernels.

2010 – Noel O'Boyle

University College Cork, Ireland. Noel was awarded the grant to both network and present his work on open source software for pharmacophore discovery and searching at the 2010 German Conference on Cheminformatics.

2009 – Laura Guasch Pamies

University Rovira & Virgili, Catalonia, Spain. Laura was awarded the Grant to do three months of research at the University of Innsbruck, Austria.

2008 – Maciej Haranczyk

University of Gdansk, Poland. Maciej was awarded the Grant to travel to Sheffield University, Sheffield, UK, for a 6-week visit for research purposes.

2007 – Rajarshi Guha

Indiana University, Bloomington, IN, USA. Rajarshi was awarded the Grant to attend the Gordon Research Conference on Computer Aided Design in August 2007.

2006 – Krisztina Boda

University of Erlangen, Erlangen, Germany. Krisztina was awarded the Grant to attend the 2006 spring National Meeting of the American Chemical Society in Atlanta, GA, USA.

2005 – Dr. Val Gillet and Professor Peter Willett

University of Sheffield, Sheffield, UK. They were awarded the Grant for student travel costs to the 2005 Chemical Structures Conference held in Noordwijkerhout, the Netherlands.
2004 – Dr. Sandra Saunders
University of Western Australia, Perth, Australia. Sandra was awarded the Grant to purchase equipment needed for her research.

2003 – Prashant S. Kharkar
Institute of Chemical Technology, University of Mumbai, Matunga, Mumbai. Prashant was awarded the Grant to attend the conference, Bioactive Discovery in the New Millennium, in Lorne, Victoria, Australia (February 2003) to present a paper, The Docking Analysis of 5-Deazapteridine Inhibitors of Mycobacterium avium complex (MAC) Dihydrofolate reductase (DHFR).

2001 – Georgios Gkoutos
Imperial College of Science, Technology and Medicine, Dept. of Chemistry. London, UK. Georgios was awarded the Grant to attend the conference, Computational Methods in Toxicology and Pharmacology Integrating Internet Resources, (CMTPI-2001) in Bordeaux, France, to present part of his work on internet-based molecular resource discovery tools.
CINF Scholarship for Scientific Excellence Co-sponsored by InfoChem and Springer

The international scholarship program of the Division of Chemical Information (CINF) of the American Chemical Society (ACS) co-sponsored by InfoChem (www.infochem.de) and Springer (www.springer.com) is designed to reward graduate and postdoctoral students in chemical information and related sciences for scientific excellence and to foster their involvement in CINF.

Up to three scholarships valued at $1,000 each will be presented at the 247th ACS National Meeting in Dallas, TX, March 16 – 20, 2014. Applicants must be enrolled at a certified college or university and must present a poster during the Welcoming Reception of the division on Sunday evening at the National Meeting. Additionally, they will have the option to show their poster at the Sci-Mix session on Monday night. Abstracts for the poster must be submitted electronically through PACS, the abstract submission system of ACS.

To apply, please inform the Chair of the selection committee, Guenter Grethe at ggrethe@att.net, that you are applying for a scholarship. Submit your abstract at http://abstracts.acs.org using your ACS ID. If you do not have an ACS ID, follow the registration instructions and submit your abstract in the CINF program in the session “CINF Scholarship for Scientific Excellence”. PACS will be open for abstract submissions on August 12, 2013, and close on October 7, 2013. Additionally, please send a 2,000-word abstract describing the work to be presented in electronic form to the Chair of the selection committee by January 31, 2013. Any questions related to applying for one of the scholarships should be directed to the same e-mail address.

Winners will be chosen based on contents, presentation and relevance of the poster and they will be announced during the Sunday reception. The contents will reflect upon the student’s work and describe research in the field of cheminformatics and related sciences.

Guenter Grethe

About InfoChem

InfoChem GmbH based in Munich, Germany, is a market leader in structure and reaction handling and retrieval. Founded in 1989, InfoChem focuses on the production and marketing of new chemical information products, including structural and reaction databases, and the development of software tools required for these applications. The main software tools provided are the InfoChem Fast Search Engine (ICFSE), the InfoChem Chemistry Cartridge for Oracle (ICCARTRIDGE) and the widely used InfoChem reaction classification algorithm CLASSIFY. InfoChem distributes one of the largest structural and reaction files worldwide, currently containing 7 million organic compounds and facts and 4 million reactions covering the chemical literature published since 1974 (SPRESI). In addition, InfoChem provides tools for the automatic recognition and extraction of chemical entities and their conversion into chemical structures as well as the semantic enrichment of chemical science documents. Springer GmbH (Berlin) has held a majority interest in InfoChem since 1991. For more information go to www.infochem.de.
Awards and Scholarship Committee Report

The 2013 Herman Skolnik Award will be presented to Dr. Richard D. (Dick) Cramer for his work in inventing the Comparative Molecular Field Analysis (CoMFA) technique and its introduction into the fields of molecular and drug design. A full-day symposium in honor of Dr. Cramer will be held on Tuesday, September 10, 2013 in the Indiana Convention Center Room 140, followed by a reception from 6:30 PM until 8:30 PM at the JW Marriott Indianapolis, Grand Ballroom 4. All CINF members are welcome to attend.

Dr. Guenter Grethe will be honored with the CINF Lifetime Award at the CINF divisional luncheon on Tuesday, September 10 for his long term membership and outstanding service and active contributions to the Division for many years. The luncheon is a ticketed event (CINF Luncheon/SE-28, Cost: $15, Time: 12:00 PM 1:30 PM, Location: Indiana Convention Center, Wabash Ballroom Section 2) and those wishing to attend should purchase a ticket with their ACS meeting registration. The CINF division has a limited number of additional luncheon tickets available for those who were not able to purchase tickets during registration; contact Andrea Twiss-Brooks, atbrooks@uchicago.edu, 773-573-8683 for more information.

The CINF Awards committee will meet briefly on Saturday afternoon, September 7. Due to the confidential nature of the committee’s work, any CINF member wishing to attend should contact the Awards chair in advance of the meeting

Andrea Twiss-Brooks  
The University of Chicago  
atbrooks@uchicago.edu

Education Committee Report

Information Skills for Graduate Students

One of the main agenda items for the CINF Education Committee at the fall Meeting will be to discuss a draft document on information skills needed for successful graduate students. Please feel free to join the Education Committee’s Meeting and participate in this discussion.

Guidelines for ACS Accredited Undergraduate Chemistry Degrees

The ACS Committee on Professional Training (CPT) is developing the next revision of the Guidelines, which will be adopted in 2014. CPT will hold an open meeting at the 246th ACS National Meeting in Indianapolis on Sunday, September 8, 2013, from noon to 1:30pm in the Indianapolis Marriott Downtown Hotel. Discussion will focus on the proposed changes to the guidelines for ACS approval. Lunch will be provided. One document under review is the Supplement on Chemical Information Skills. The CINF Education Committee
plans to send comments to CPT about proposed changes to the Chemical Information Searching part of the Guidelines. The Education Committee also plans to review the Chemical Information Literacy Competencies for Undergraduate Chemistry Students to see if any revisions are needed.

Explore Chemical Information Teaching Resources (XCITR)

FIZ CHEMIE Berlin was very generous to develop and host the XCITR website, but due to changing circumstances is no longer able to do so. The Royal Society of Chemistry (RSC) has graciously agreed to host XCITR in 2013. The URL has stayed the same: http://www.xcitr.org.

At the end of the year, RSC and the XCITR Board will evaluate the future support needed for XCITR. There are approximately 50 items in the XCITR collection now. The current goal is to encourage more submissions from busy colleagues so that content they have already created can be used by other librarians and by chemistry instructors. The CINF Education Committee has initiated efforts to increase the number of items in XCITR with the goal of doubling the size of the collection by the end of 2013.

Submissions of educational materials from information providers, chemistry instructors, and librarians are desired. Submissions can be descriptions with links to web pages or videos on a provider’s website or deposition of documents. Because the audience is broad, XCITR is an excellent way to share information with colleagues and clients.

Biennial Conference on Chemical Education 2014.

The Biennial Conference on Chemical Education (BCCE) is a national meeting sponsored by the Division of Chemical Education (DivCHED) of the American Chemical Society (ACS). The 23rd BCCE will be held at Grand Valley State University, Michigan, on August 3-7, 2014. Please see http://www.bcce2014.org/. The theme for the 2014 meeting is Greener on the Grand: Empowering Chemical educators for a Greener Tomorrow.

Call for workshops begins June 3, 2013 and the call for symposia begins August 1, 2013. The deadline for submitting proposals for workshops and symposia is December 6, 2013. Proposals for symposia that fit into tracks of: Education Research, Advances in Pedagogy, Computers and Technology in Chemistry, Green Chemistry, Chemical Informatics, Global Chemistry Education, Chemistry for the Public, Discipline Specific Symposia, and any other topics you wish to propose will be welcome.

At the Fall ACS National Meeting, the CINF Education Committee will decide whether or not to submit a proposal for BCCE 2014. Any CINF members interested in participating in BCCE 2014 should contact Grace Baysinger.

Grace Baysinger
Stanford University
graceb@stanford.edu

Chemical Information Bulletin, 2013, 65 (3)
Nominating Committee Report

The results of the election are in and the newly elected are:

2014 Chair Elect - Rachelle Bienstock
2014-2015 Secretary - Leah McEwen

Congratulations!

Sue Cardinal
CINF Teller
Interview with Maureen Rouhi, Editor-in-Chief of Chemical & Engineering News

A. Maureen Rouhi is the Editor-in-chief of Chemical & Engineering News, the weekly news magazine of the American Chemical Society (ACS). It reports current events in the chemistry enterprise, including recent advances in research, education, industry, funding, and regulatory policy. C&EN has a weekly circulation of about 110,000. It is produced by staff based in Washington, D.C.; Edison, NJ; Boston, MA; Chicago, IL; Dallas, TX; Houston, TX; Oakland, CA; Berlin; Hong Kong; and London. Rouhi is responsible for about 50 full-time staff. Rouhi received B.S. and M.S. degrees in agricultural chemistry from the University of the Philippines, Los Baños, and a Ph.D. from the University of London. Her professional experience includes independent research and teaching in the Philippines and in Iran, scholarly publishing, journalism, and news organization management. She began her scholarly publications career as a copy editor trainee at the American Society for Microbiology. In 1987, she joined the ACS, where she has worked ever since, except for an 11-month detour to the American Pharmaceutical Association (now the American Pharmacists Association), where she served as managing editor of the Journal of Pharmaceutical Sciences.

At ACS, she worked in books production, books acquisitions, and production of multimedia continuing-education courses. Her journalism career began when she joined C&EN in 1994. The reporting assignments that she liked most were those about chemical ecology, the chemical senses, natural products, organic chemistry, pharmaceuticals, fine chemicals, and the history of chemistry. From 2004 to 2012, she was deputy editor-in-chief of C&EN. In that capacity, she focused on the daily operations that help make C&EN every chemist's first choice for news about the chemical enterprise. She took over as editor-in-chief in September 2012. In this capacity, she aims to take C&EN to the next level of excellence and achievement by expanding its influence, extending its global reach, and strengthening its digital presence. Rouhi’s leadership and management skills have been formed by 24 years of service at the ACS. The world’s largest scientific society, ACS is a nonprofit professional organization with a multidisciplinary membership of more than 163,000 chemists and chemical engineers. Its revenues come primarily from sales of more than 40 scientific journals and information services provided by Chemical Abstracts Service (CAS), the biggest and most comprehensive collection of chemical information in the world.
Svetla Baykoucheva:

Maureen, congratulations on your appointment as Editor-in-Chief of the ACS flagship magazine, Chemical & Engineering News. Your educational and professional backgrounds are impressive, and your life story is unique. What were the most important events and circumstances that have influenced the choices you have made and how have they shaped your career?

Maureen Rouhi:

Thank you, Svetla, for the invitation to talk to you about C&EN. My personal story affirms how little control we have in shaping the way our lives unfold. As a college student in the Philippines, I planned an uncomplicated trajectory: graduate school and academia. The Philippine government awarded me a full scholarship to study for a Ph.D. in the U.K. In return I would return to the country and serve in the university that nominated me for the scholarship. All was going as planned until I met my future husband, who is not from the Philippines but from Iran. Marriage changed the direction of my life. After living in Iran for four years, we emmigrated to the U.S. Needing to find a job as quickly as possible, I grabbed the chance to start a career in science publishing. That was a lucky choice, because it is what brought me to the ACS and eventually to C&EN.

SB:

You are inheriting a magazine that has been in existence for almost 90 years. Your immediate predecessors—Rudy Baum, who recently retired, and Madeline Jacobs, who stepped down as Editor-In-Chief in 2004 and now leads the Society as the Executive Director & CEO, have left a memorable legacy. The way scientists perform research and report their findings is rapidly changing. Science has become very interdisciplinary. How will this impact the directions to which you will be taking the publication? Is there going to be a change in the name of the magazine?

MR:

Indeed, we are celebrating C&EN's 90-year anniversary all year this year. We have a lot to celebrate, thanks to the efforts, leadership, and vision of previous editors-in-chief. Observant readers know that once a month a C&EN editor's page is devoted to revisiting stories in C&EN Archives. Weekly, our Tumblr blog, The Watch Glass, offers stunning images and excerpts taken at random from C&EN Archives. And we will have celebratory events at the ACS National Meeting in Indianapolis. In the morning of Tuesday, Sept. 10, we are hosting a free webinar titled Food Fraud: How Scientists Detect It & What You Should Know. In the evening of the same day, we are hosting celebrity chef and author Alton Brown in a free live event and a reception for attendees after the show. I’m looking forward to seeing you and your readers at these festivities.

C&EN's vision is to be the first choice for reliable news and information about the chemical enterprise. We report science as it happens, including the influences that affect its practice, such as globalization, collaboration, technology, R&D support, government policies & regulations, emerging fields of inquiry, education, etc.
Changing the magazine’s name is a topic that comes up periodically. C&EN enjoys great name recognition, and we are not considering a name change at this time.

SB:

For many years, you were a “lab-bench” reporter writing about science. I remember, from my days working at the ACS Library & Information Center, how meticulous you were in researching the topics you were writing about. How will your background and work as a PhD chemist and a reporter influence your leadership at C&EN?

MR:

One of my biggest tasks as leader of C&EN is being the readers’ advocate. Communicating the stories of scientists who do great work can be a lot of fun when the work is in areas with intrinsic mass appeal. Work related to food and health, for example, is relatively easy to make accessible to a wide audience. The same cannot be said for a lot of other important work. For example, discoveries about new modes of chemical bonding or the physics behind instrumental advances are difficult subjects, yet critical in advancing chemistry. My role is to make sure that even the complex topics are understandable to nonexperts. Having been both a researcher and a reporter, I am all too aware of the gap between the world of scientists and the needs of the general public. Making sure that C&EN stories bridge that gap is part of what I do every day at C&EN.

SB:

Scientific publishing has changed significantly, and the Open Access movement has forced commercial and other publishers to change their policies and become more “user-friendly.” How do you see the STEM publishing field changing in the next few years?

MR:

Open access has been gaining a lot of momentum recently, especially in Europe. In the U.S., the Obama Administration is pushing for greater access to federally-funded research, and U.S. science publishers are under pressure to adapt to a new scholarly publishing landscape.

SB:

How are the issues of C&EN being planned? Is there a general view of what the future issues will be covering—and who is making these decisions? How much independence do the individual reporters and editors have in making their choices?

MR:

Planning in C&EN is very much a collaborative process. Daily, C&EN reporters are communicating with editors about potential stories for news sections. Weekly and monthly, C&EN’s senior staff and leadership go over a long list of story ideas to figure out how they would be developed and when they would be ready to go in the magazine or online. Annually, C&EN editorial, advertising, and marketing staff meet to plan the next year’s editorial calendar. C&EN reporters have a lot of flexibility in selecting what stories
to pursue, but their story ideas are always rigorously vetted by editors and others—for example, the online editor and the design director—who may have input in how a story is developed.

SB: How do reporters and editors at C&EN gather information? What resources are they using to do that?

MR: A lot has changed in how journalists do research since the birth of the Internet. The Internet is great and convenient, as long as reporters know where to go for reliable information. Reporting through social media, especially Twitter, is a recent phenomenon that some C&EN reporters are using to great advantage. Of course, C&EN reporters still gather information through all the traditional ways: by reading documents, attending events, and talking to people.

SB: How does C&EN preserve its objectivity and neutrality when publishing articles on topics that may be perceived as controversial—from the point of view of ACS or other organizations, or even by the government?

MR: As a news organization, C&EN is committed to the highest principles of excellence in journalism. When covering controversial topics, we go to great lengths in our research to ensure that we probe as many perspectives as possible. We aim to be neutral and objective, but it’s possible that a story could emphasize one perspective more than others. That’s because not all perspectives have equal weight, and part of our role as journalists is to consider the validity of information we gather and to come to conclusions on the basis of facts and arguments that, in our judgment, really matter in clarifying or resolving the controversy.

SB: If you could imagine, for a moment, that you are still a science journalist, what is the topic you would be most interested to write about right now?

MR: I do still consider myself a science journalist, except that I do not have time anymore to develop the story ideas I come across. I was at Case Western Reserve University in early June to participate in the celebration of the 50-year anniversary of the Department of Macromolecular Science & Engineering. I came back from that conference with several story ideas about advanced materials. This field is something I did not write much about when I was a full-time reporter, but I was really excited about what I heard at Case Western Reserve. I’ve given those story ideas to various staff, and sooner or later, the stories will appear in C&EN.
SB:

Could you tell us something about yourself? What interests you, beyond your professional life? What books do you like to read? What topics are you following right now? What do you like to do when not working?

MR:

I’m a great believer in continuous learning, and for many years now, I’ve tried to learn something new every year. For example, I’ve taken acting classes for business people, offered by the Shakespeare Theatre in Washington, DC. I also enrolled in ballroom dancing classes until I had to stop because of problems with my left shoulder and right knee. I’ve been practicing yoga for a long time, for both physical and mental fitness. And with luck, this fall, I hope to enroll in a class on improvisational acting.

By Svetla Baykoucheva
University of Maryland College Park
sbaykouc@umd.edu
Book Reviews


This familiar essential reference has appeared in its 15th edition. Several changes have been made and the first to be noticed is that although it’s still The Merck Index, it’s now published by the Royal Society of Chemistry (RSC). The previous Merck Index staff prepared this edition but publication and preparation will henceforth be by the RSC. As previously with the 14th edition, a purchase also provides a one-year complimentary (i.e. free) subscription to the web version ( > 11,500 monographs), which will be updated quarterly. However, there is no CD provided for the 15th even though the price has been increased by $25.

As for the content, there are > 10,000 monographs covering 18,000 substances (some monographs describe additional related substances) and > 3500 monographs have been updated. To the approximately two dozen Supplementary Tables have been added non-proprietary pharmaceutical name stems which should prove very useful to consumers. An increase of 50 brings the total number of named organic reactions to > 500. Cross indexes include Name, CAS Registry Number (CAS RN), chemical formulas, and biologic or therapeutic activity. Monographs list the Merck Index name and, depending on availability, synonyms, trade names, CAS RN, molecular weight and formula, elemental analysis, toxicity data or references, and preparation references including patent numbers.

The web version can be searched by using two modes: quick and primary. Quick searches can be performed on names, molecular formulas, or molecular weights. Primary searches are more advanced and also include searches of almost any data or information in the database. Either exact or range searches can be performed for data elements. Structure of searching is available, by exact, substructure, or similarity.

Purchase of a print copy is highly recommended for a wide range of students, teachers and researchers. In addition to print copies, libraries should consider acquiring access to the online files. For pricing, contact the publisher.


Continuing a practice made from at least the 5th and 6th Editions onward, Appendix A of the 7th Edition of this massive text describes the literature of organic chemistry. Both the print and electronic primary and secondary literature are covered. Primary sources include both journals and patents. Secondary sources include titles listings (Current Contents, etc.), abstracts (especially CA), Beilstein, and data tables (Heilbron, Handbook of Chemistry and Physics, Merck Index, etc. Several categories of reviews are covered. Literature searching
is covered extensively, both by use of printed and online sources. Included in the latter are online searching in general including STN, SciFinder, *Science Citation Index* (SCI), journal articles (CASSI et al.), and Reaxys.

Overall, this Appendix is a fair primer, at least for print sources, and has been updated from the 6th Edition. However much of the online coverage is erratic and even erroneous. The recommended method for searching printed *Beilstein* is described. Gary Wiggins’ book, *Chemical Information Sources* (but not the Wikibook successor), is finally cited as describing both print and online searching. Most of the links for the CAS Registry System are no longer active and the Material Safety Data Sheets (MSDS) file cited is no longer available. SciFinder and SciFinder Scholar are no longer separate entities. The first two paragraphs of the SciFinder section instead describe STN. The online section on SCI instead mostly describes the print version and SCI is erroneously described as not available on STN. The emphasis on literature searching in a basic text is to be admired, but better coverage and updating is required for future editions.

While I was at it, I briefly reviewed the rest of the text. The book is extensively referenced and coverage of the literature is quite current. More than 20,000 references are cited, 5500 of them new. Chapters 1-5 cover organic chemical structure, 6-9 cover mechanisms, and 10-19 cover reactions. Aimed at the graduate course level, it’s also recommended for an advanced undergraduate course. The emphasis on reactions should be an aid to synthesis planning. Admittedly several specialized areas of organic chemistry (terpenes, steroids, etc.) are covered very little or not at all, but students should have enough experience to find the appropriate coverage in books or reviews.

This book will be a welcome addition to my organic chemistry bookshelf (very dated). A big improvement compared to my graduate school days when one had to consult several additional texts to get this advanced material. I do intend to initiate some corrective measures with the author, per his invitation.

Bob Buntrock
Orono, ME
buntrock16@roadrunner.com
Literature Digest


Song Yu
Columbia University
sy2133@columbia.edu
At both the individual and professional levels, we are dealing with a massive flow of information every day. Data are growing faster than the means and the wherewithal to store, transmit, analyze, manage and support them, leaving a widening “data deluge gap.” In today’s data-driven world, information has enormous value when it is analyzed, shared and used. It is this need that is aimed to transform the World-Wide Web to a richer web, the “Semantic Web” or the “Linked data Web” where a common framework allows applications to share and reuse data. Semantic web is about “meaningful computing”, making links to connect related data. The idea of semantic technology is to define, link, integrate, and analyze data from various data sets to obtain new information stream so as to help in more effective discovery, automation, integration, and reuse across various applications.  

Currently, web-based information is based primarily on documents written in HTML that allows visual presentation but provides limited avenues of describing the content. Semantic Web technology aims to address this shortcoming, using RDF, OWL and XML languages. These machine-readable languages allow intelligent software to be written while automating the analysis. Semantic technology incorporates morphological, logical, grammatical and natural language analysis while integrating both structured and unstructured data, thus not only translating into higher precision and recall when searching for information, but also delivering all of the most important and accurate data. 

The Semantic Web did not evolve overnight and decades of research work are in place since the appearance of the first paper in 1980. Different types of articles from scholarly publications through proceeding papers, editorial material and reviews run the gamut of subjects from computer science, through drug discovery, literature, social science and medicine (see references 6 and 7 for examples). In addition, for some recent information on “semantification” of chemistry see the winter 2012 CIB.

Tracking the growth of the “Web of Data” over the last three decades using citation and data from “Web of Science” provides interesting results. Data was collected for fixed time periods and analyzed over a set of parameters including areas of research, language, and country of research. Though computer science and engineering have always been prominent fields of study, there is the slight positive shift in the contributions from Biochemistry-Molecular Biology and Health care sciences (data not shown). It is worth noting the increase in contributions from China over a period of time (see tables). I believe the research community has just taken the first few steps. Though the Semantic web’s future looks promising in transforming the Web of static data to a global computational resource, we still have miles to go.
Tables 1 and 2: Analysis of selected time-frame records based on countries. The time-based records are generated from searching for "semantic* web" keyword in the Web of Science:

**Table 1. 1990–1999**

<table>
<thead>
<tr>
<th>Countries/Territories</th>
<th>Records</th>
<th>% of 196</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA</td>
<td>99</td>
<td>50.5</td>
</tr>
<tr>
<td>France</td>
<td>14</td>
<td>7.1</td>
</tr>
<tr>
<td>Germany</td>
<td>11</td>
<td>5.6</td>
</tr>
<tr>
<td>Italy</td>
<td>9</td>
<td>4.6</td>
</tr>
<tr>
<td>Australia</td>
<td>7</td>
<td>3.6</td>
</tr>
<tr>
<td>England</td>
<td>7</td>
<td>3.6</td>
</tr>
</tbody>
</table>

**Table 2. 2000–2004**

<table>
<thead>
<tr>
<th>Countries/Territories</th>
<th>Records</th>
<th>% of 2235</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA</td>
<td>576</td>
<td>25.8</td>
</tr>
<tr>
<td>China</td>
<td>217</td>
<td>9.7</td>
</tr>
<tr>
<td>Germany</td>
<td>204</td>
<td>9.1</td>
</tr>
<tr>
<td>England</td>
<td>185</td>
<td>8.3</td>
</tr>
<tr>
<td>Italy</td>
<td>143</td>
<td>6.4</td>
</tr>
<tr>
<td>France</td>
<td>104</td>
<td>4.7</td>
</tr>
</tbody>
</table>

Notes: Only the first six major contributors are shown.
Tables 3 and 4: Analysis of selected time-frame records based on countries. The time-based records are generated from searching for "semantic* web" keyword in the Web of Science:

Table 1. 2005–2009

<table>
<thead>
<tr>
<th>Countries/Territories</th>
<th>Records</th>
<th>% of 9036</th>
</tr>
</thead>
<tbody>
<tr>
<td>China</td>
<td>1653</td>
<td>18.3</td>
</tr>
<tr>
<td>USA</td>
<td>1237</td>
<td>13.7</td>
</tr>
<tr>
<td>Germany</td>
<td>753</td>
<td>8.3</td>
</tr>
<tr>
<td>England</td>
<td>653</td>
<td>7.2</td>
</tr>
<tr>
<td>Italy</td>
<td>599</td>
<td>6.6</td>
</tr>
<tr>
<td>Spain</td>
<td>430</td>
<td>4.8</td>
</tr>
</tbody>
</table>

Table 1. 2010–present

<table>
<thead>
<tr>
<th>Countries/Territories</th>
<th>Records</th>
<th>% of 4635</th>
</tr>
</thead>
<tbody>
<tr>
<td>China</td>
<td>761</td>
<td>16.4</td>
</tr>
<tr>
<td>USA</td>
<td>687</td>
<td>14.8</td>
</tr>
<tr>
<td>Spain</td>
<td>420</td>
<td>9.1</td>
</tr>
<tr>
<td>Germany</td>
<td>387</td>
<td>8.4</td>
</tr>
<tr>
<td>England</td>
<td>343</td>
<td>7.4</td>
</tr>
<tr>
<td>Italy</td>
<td>329</td>
<td>7.1</td>
</tr>
</tbody>
</table>

Notes: Only the first six major contributors are shown.
If the semantic web sounds interesting to you, be sure to check out some of the upcoming talks at the Fall 2013 ACS meeting. Here are a few that discuss the semantic web:

1:35 – 57 – Building support for the semantic web for chemistry at the Royal Society of Chemistry. Valery Tkachenko, tkachenkov@rsc.org, Colin Batchelor, Jon Steele, Alexey Pshenichnov, Antony J. Williams.


3:20 – 60 – Practical semantics in the pharmaceutical industry: The Open PHACTS project. Antony J. Williams, williamsa@rsc.org.

3:50 – 61 – Enabling the translational medicine and drug discovery information workflow. David Evans, david.evans@reedelsevier.ch, Timothy Hoctor, Jacqui Mason, Pieder Caduff.

4:50 – 63 – Application of text mining and semantic technology in external intelligence surveillance. Yiqun Li, lihy@lilly.com.

References:


The American Chemical Society Division of Chemical Information is very fortunate to receive generous financial support from our sponsors to maintain the high quality of the Division’s programming and to promote communication between members at social functions at the ACS Fall 2013 National Meeting in Indianapolis, IN, and to support other divisional activities during the year, including scholarships to graduate students in Chemical Information. The Division gratefully acknowledges contribution from the following sponsors:

**Gold**
- Certara

**Silver**
- ACS Publications
- Bio-Rad Laboratories
- Royal Society of Chemistry

**Bronze**
- Accelrys
- ACS Symposium Series
- InfoChem
- Elsevier
- Journal of Cheminformatics
- Optibrium
- PerkinElmer
- Thieme Chemistry
- Novartis
- Kilmorie Clarke

**Contributors**
- PID Analyzers
- RSC eScience

Opportunities are available to sponsor Division of Chemical Information events, speakers, and material. Our sponsors are acknowledged on the CINF web site, in the Chemical Information Bulletin, on printed meeting materials, and at any events for which we use your contribution. For more information please review the Sponsorship Brochure at [http://www.acscinf.org/PDF/Fundraising_Brochure.pdf](http://www.acscinf.org/PDF/Fundraising_Brochure.pdf). Please feel free to contact me if you would like more information about supporting CINF.

Rajarshi Guha, Chair, Fundraising Committee
Email: Fundraising@acscinf.org
Tel: 814-404-5449

The ACS CINF Division is a non-profit tax-exempt organization with taxpayer ID no. 52-6054220.
Bio-Rad — Access Over 1.4 Million Spectra Today at Your University

It is with great pleasure that Bio-Rad Laboratories, a long-time sponsor of CINF, invites your school to participate in a free 30-day evaluation trial of KnowItAll U—the world’s largest collection of spectral data.

This resource has been used successfully in the teaching curriculum, academic research in chemistry, biology, & other applied science disciplines, as well as in a library setting.

- KnowItAll U provides every student and faculty member unlimited campus-wide access to 1.4 million spectra (NMR, IR, Raman, MS, UV-Vis).
- Access via the KnowItAll AnyWare web browser interface or through the KnowItAll software—a full suite of tools to search, manage, analyze spectra; predict NMR spectra; draw structures; & more!

With this trial, your school will have complete access to the spectra and software for 30 days.

Please contact Farrel Borden at 267-322-6938 or register online at www.knowitallu.com/trial. It’s that easy!

For more information, you can also visit www.knowitallu.com or view a recording of our “Introduction to KnowItAll U” webinar at www.knowitall.com/webex_archives (presented by Farrel Borden).
Spanning the gap between discoveries published in patents and journals, the peer-reviewed e-books of the ACS Symposium Series contain essential research by the world’s leading scientists spanning all disciplines and applications. The series provides more than 1,300 books, including the entire Advances in Chemistry Series, and it is fully searchable by keyword, title, author, year of publication, and division sponsor.

- **A Format Readers Prefer**
  Hosted on the award-winning ACS web platform, each chapter is formatted like a journal article and available as a high-resolution PDF viewable on your favorite device.

- **Multidisciplinary Content**
  Supports research in areas such as environmental chemistry, toxicology, agriculture, pharmaceuticals, polymer chemistry, materials, and more.

- **Top Authors**
  Features the work of 32 Nobel Laureates, including Melvin Calvin, William N. Lipscomb, Henry Taube, George A. Olah, Alan J. Heeger, and Herbert C. Brown.

Find out more at pubs.acs.org/series/symposium.
InfoChem — SPRESIweb version 2.13 successfully released

InfoChem is pleased to announce the successful release of SPRESIweb 2.13. The well-known web-based application gives direct access to the complete content of the SPRESI structure and reaction database that includes now 5.62 million compounds and 4.34 million reactions, covering the literature for the time period 1974-2012. Click here or visit www.spresi.com for more information about SPRESIweb and to apply for a test account.

Name Reaction search to be included in SPRESImobile App

InfoChem is pleased to announce a new release of SPRESImobile in September 2013. The free App, successfully introduced on iTunes in March 2012 and developed to perform structure and reaction searches on mobile devices, gives access to a subset of the SPRESI reaction data (ChemReact).

Name Reaction searching will be integrated in the new iPhone and iPad App version: after registration users will be able to see a new button for name reaction searches in the reaction search form. Furthermore, all users (also not registered ones) will see a new link in the reaction hit list, when named reactions are hits for a reaction search. For more information about SPRESImobile please click here.

Please feel free to contact us for more information about InfoChem, our current research projects and our products.
PerkinElmer Offers White Paper on Mobile Technology for the Scientific Workflow

PerkinElmer, Inc., a global leader focused on improving the health and safety of people and the environment, offers a complimentary white paper on the advantages of using mobile technology in the scientific workflow. The paper, “Liberating Today’s Scientists with App-Enabled Mobile Devices,” argues that apps in science can increase productivity in and out of the lab and foster the collaboration that sparks innovation by providing rapid anywhere, anytime access to data and enabling remote generation of data and experiment control.

Having already revolutionized consumer experiences and infiltrated the business world, mobile technology is making inroads into the lab. Scientific apps are gradually appearing on smartphones and tablets. Leading companies that embrace mobility and find ways to encourage its use in the lab will help scientists improve their productivity and accelerate discoveries.

PerkinElmer, envisioning an environment where scientists perform defined tasks and generate data using mobile devices, recently introduced its first apps for science. ChemDraw and Chem3D for iPad are new chemical structure drawing and visualization apps for the iPad that empower scientists and students to collaborate outside of the lab and classroom. Mobile extensions of its gold-standard ChemDraw software suite and Chem3D molecular viewing solution, PerkinElmer’s first apps are available from the Apple App Store.

The white paper covers the exponential rise in the use of apps for consumers and in business, how their introduction in the lab can benefit scientists, PerkinElmer’s strategy to develop its first apps, and details on how ChemDraw and Chem3D for iPad function. For a free download of the paper, visit:

Thieme Chemistry — SYNTHESIS and SYNLETT experimental procedures will soon be available in SciFinder

Chemical Abstracts Service (CAS) and Thieme Chemistry have agreed to make hundreds of thousands of experimental procedures from SYNTHESIS and SYNLETT articles available in the database SciFinder. This will give chemists all over the world easy access to the research in Thieme Chemistry’s synthetic organic chemistry journals.

Technical Program

ACS Chemical Information Division (CINF)
246th ACS National Meeting and Exposition

Sept. 8–12, 2013, Indianapolis, IN
J. Garritano, Program Chair

Sunday, September 8, 2013

Current Challenges in Cheminformatics: Exploiting Information and Knowledge in Structured and Unstructured Environments – AM Session
Cheminformatics Tools and Methods

Indiana Convention Center
Room: 140
Cosponsored by COMP

Neil Kirby, Dirk Tomandl, Organizers
Dirk Tomandl, Presiding

8:10 am – 12:00 pm

8:10 – Introductory Remarks

8:15 – 1 – New insights and tools for capturing, validating, and utilizing structure/property data: Curiously similar approaches for drug discovery and materials science. Curt M Breneman, brenec@rpi.edu, Ke Wu, Lisa Morkowchuk, Jed Zaretzki, Sourav Das, Michael Krein.

8:45 – 2 – Scaffold-based reasoning approaches for cross-assay structure-activity relationship identification. Christos A Nicolaou, c.nicolaou@lilly.com, Quan Liao, Suntara Cahya, Jibo Wang.


9:45 – Intermission

9:55 – 4 – Computational approach of thermodynamic fragmentation applied to formulating badly soluble actives. Johannes Fraaije, j.fraaije@chem.leidenuniv.nl, Monica Bulacu, Ruben Gracia, Shyamal K Nath.


11:25 – 7 – Plexus: A clean web application for structural data analysis David Deng, ddeng@chemaxon.com, Andras Stracz.

11:55 – Concluding Remarks
Sunday, September 8, 2013

Chemistry on Tablet Computers – AM Session

Indiana Convention Center
Room: 141

David Martinsen, Martin Braendle, Organizers
Martin Braendle, Presiding

8:05 am – 12:00 pm

8:05 – Introductory Remarks

8:10 – 8 – Apps and approaches to mobilizing chemistry from the Royal Society of Chemistry. Antony J. Williams, williamsa@rsc.org, Valery Tkachenko, Dmitry Ivanov, Will Russell.

8:40 – 9 – ChemDraw, iPads, and collaboration tools in the classroom: Results of a joint PerkinElmer and McGraw Hill pilot at the organic chemistry undergraduate level. Robin Y Smith, robin.smith@perkinelmer.com, Hans C Keil, hans.keil@perkinelmer.com, Tamara Hodge.

9:10 – 10 – Can I get that to go? Reading research articles on a tablet. Jeff Lang, j_lang@acs.org.

9:40 – Intermission

9:55 – 11 – JSmol: Full-service molecular visualization on the Web without Java. Robert Hanson, hansonr@stolaf.edu, Zhou Renjian, Takanori Nakane, Paul Pillot.

10:25 – 12 – Enabling chemistry on-the-go with modern solutions. Tamsin E Mansley, tamsin.mansley@dotmatics.com; Graeme E Dennis, Shikha O’Brien.

10:55 – 13 – Tablets in the lab: Enabling the flow of chemical synthesis data into a chemistry repository. Simon J Coles, s.j.coles@soton.ac.uk, Richard J Whitby, Aileen Day, Cerys Willoughby, Valery Tkachenko, Jeremy G Frey, Antony J Williams

11:25 – 14 – New strategy to engage mobile computing users and developers. Steven M Muskal, smuskal@eidogen-sertanty.com.

11:55 – Concluding Remarks

Sunday, September 8, 2013

Current Challenges in Cheminformatics: Exploiting Information and Knowledge in Structured and Unstructured Environments – PM Session

Data Access, Usage and Pitfalls

Indiana Convention Center
Room: 140

Cosponsored by COMP

Dirk Tomandl, Neil Kirby, Organizers
Dirk Tomandl, Presiding

1:10 pm – 3:50 pm

1:10 – Introductory Remarks

1:15 – 15 – Improving access to data: A distributed approach. Graeme E Dennis, graeme.dennis@dotmatics.com, Tamsin E Mansley, Shikha O’Brien.

1:45 – 16 – Digital fractionation: Using NoSQL technology to extract biochemical data from big data repositories. Lewis Y Geer, lewis.geer@nih.gov, Lianyi Han, Yanli Wang, Evan E Bolton, Siqian He, Bo Yu.
2:15 – Intermission

2:20 – 17 – Exploring interconnectivity of molecular biology databases and data mining tools. Stefan M Furrer, stfurrer@indiana.edu, David J Wild.

2:50 – 18 – Caveats in the use of pharmaceutical drug discovery data. Terry R Stouch, tstouch@gmail.com.

3:20 – 19 – Extracting synthetic knowledge from reaction databases. Orr Ravitz, ravitz@simbiosys.com, James Law, Anthony Cook, A. Peter Johnson.

4:00 – 20 – Validation and characterization of chemical structures derived from names and images in scientific documents. John B Kinney, john.b.kinney@dupont.com.

4:30 – 21 – Tackling the difficult areas of chemical entity extraction: Misspelt chemical names and unconventional entities. Daniel M Lowe, daniel@nextmovesoftware.com, Roger A Sayle.

5:00 – 22 – OPSIN: Taming the jungle of IUPAC chemical nomenclature Daniel M Lowe, daniel@nextmovesoftware.com, Peter Murray-Rust, Robert C Glen.

Sunday, September 8, 2013

Graduate Student Research Symposium in Cheminformatics, Information Science, and Library Science – PM Session

Indiana Convention Center
Room: 141

Cosponsored by COMP

Gary Wiggins, Organizer
Sunday, September 8, 2013

CINF Scholarship for Scientific Excellence – EVE Session

Indiana Convention Center
Room: Wabash Birm 1/2

Guenter Grethe, Organizer

6:30 pm – 8:30 pm

29 – Enhanced ranking of PknB inhibitors using data fusion methods. Abhik Seal, abseal@indiana.edu, David John Wild.


31 – Harvard Clean Energy Project: From big data and cheminformatics to the rational design of molecular OPV materials. Johannes Hachmann, jh@chemistry.harvard.edu, Roberto Olivares-Amaya, Alan Aspuru-Guzik.

Monday, September 9, 2013

Integrative Chemogenomics Knowledge Mining Using NIH Open Access Resources – AM Session

Indiana Convention Center
Room: 140

Rajarshi Guha, Tudor Oprea, Paul Clemons, Organizers
Rajarshi Guha, Presiding

8:50 am – 12:00 pm

8:50 – Introductory Remarks

8:55 – 32 – Pushing chemical biology data through the pipes: Architecting and extending the BARD API. Rajarshi Guha, rajarshi.guha@gmail.com, John Braisted, Ajit Jadhav, Dac–Trung Nyguen, Tyler Peryea, Noel Southall.

9:20 – 33 – QSAR modeling on the web. Diane Pozefsky, pozefsky@cs.unc.edu, Diptorup Deb, Chi Xie, Alexander Sedykh, Alex Tropsha.


10:10 – Intermission


10:45 – 36 – BADAPPLE promiscuity plugin for BARD: Evidence-based promiscuity scores. Jeremy J Yang, jeremyjyang@gmail.com, Oleg Ursu, Cristian G Bologa, Anna Waller, Larry A Sklar, Tudor I Oprea.

11:30 – Panel Discussion

Monday, September 9, 2013

Role and Value of Social Networking in Advancing the Chemical Sciences – AM
Session Social Media for the Individual Scientist and to Support Education

Indiana Convention Center
Room: 141

Cosponsored by CHED, SCHB, YCC

Antony Williams, Jennifer Maclachlan, Organizers
Jennifer Maclachlan, Presiding

8:15 am – 12:00 pm

8:15 – Introductory Remarks

8:20 – 37 – @ChemConnector and my personal experiences in participating in the expanding social networks for science. Antony J. Williams, williamsa@rsc.org.

8:45 – 38 – Get to the point! Techniques for organizing your online social networking skills. Tom Ruginis, happilabs@gmail.com.


9:35 – Intermission

9:45 – 40 – Using social tools collaboratively to communicate and advance science. Christopher McCarthy, c_mccarthy@acs.org, Christine Brennan-Schmidt, c_schmidt@acs.org.

10:10 – 41 – Social media in cheminformatics education. David J Wild, djwild@indiana.edu, Robert Belford.

10:35 – 42 – ChemConf to ConfChem: Twenty years and counting. Robert E. Belford, rebelford@ualr.edu, Harry E. Pence.

11:00 – Intermission

11:10 – 43 – Small chemical businesses and the importance of being social. Jennifer L Maclachlan, pidgirl@gmail.com.

11:35 – 44 – XCITR (Explore Chemical Information Teaching Resources): A community resource for chemistry instructors, chemistry librarians, and chemical information providers. Grace Baysinger, graceb@stanford.edu, Andrea Twiss-Brooks, Guenter Grethe, Gregor Fels.

Monday, September 9, 2013

Science-Based Policy Development in the Environment, Food, Health, and Transport Sectors – PM Session

Indiana Convention Center
Room: 140

Cosponsored by AGFD, ANYL, ENVR, MEDI

William Town, Organizer
William Town, Presiding

1:15 pm – 5:05 pm

1:15 – Introductory Remarks

1:20 – 45 – Adaptive management tools for engineered nanomaterials in municipal wastewater effluents. Thomas A. Duster, tduster@nd.edu.

1:50 – 46 – Role of STEM data and information in an environmental decision-making scenario: The case of climate change. Frederick W Stoss, fstoss@buffalo.edu.

2:20 – 47 – Identification of pathways of
toxicity to predict human effects. Helena T Hogberg, hhogberg@jhsph.edu, Thomas Hartung.

2:50 – Intermission

3:05 – 48 – Role of education and training in supporting science-based policy development. Rodger D Curren, rcurr@iivs.org, Hans A Raabe, Brian C Jones.


4:05 – 50 – Role of regulatory science in reducing the public health impact of tobacco use. Christopher J Proctor, christopher_proctor@bat.com, Chuan Liu.

4:35 – 51 – Systematic and structural risk analysis approaches for establishing maximum levels of essential nutrients and other bioactive substances in fortified foods and food supplements. David P Richardson, info@dprnutrition.com.

Monday, September 9, 2013

Role and Value of Social Networking in Advancing the Chemical Sciences – PM Session Social Media to Share Science with the Community and the Sharing of Chemical Information

Indiana Convention Center
Room: 141

Cosponsored by CHED, SCHB, YCC

Antony Williams, Jennifer Maclachlan, Organizers
Antony Williams, Presiding

1:00 pm – 3:45 pm

1:00 – 52 – Exploiting the digital landscape to advance the chemical sciences. Bibiana Campos-Seijo, camposseijo@rsc.org, Will Russell, Vibhuti Patel.

1:25 – 53 – Using social media as a way to communicate science to the community. George Ruger, gruger04@yahoo.com.

1:50 – 54 – Collaborating to convey chemical knowledge through Wikipedia. Martin A Walker, walkerma@potsdam.edu.

2:15 – Intermission

2:25 – 55 – Tools and strategies for sharing chemical information and research on the web. Jean-Claude Bradley, bradlejc@drexel.edu, Andrew Lang.


3:15 – Panel Discussion

3:40 – Concluding Remarks

Monday, September 9, 2013


Indiana Convention Center
Room: 142

David Wild, Jan Kuras, Organizers
David Wild, Jan Kuras, Presiding

1:30 pm – 5:25 pm

1:30 – Introductory Remarks
1:35 – 57 – Building support for the semantic web for chemistry at the Royal Society of Chemistry. Valery Tkachenko, tkachenkov@rsc.org, Colin Batchelor, Jon Steele, Alexey Pshenichnov, Antony J. Williams.


3:05 – Intermission

3:20 – 60 – Practical semantics in the pharmaceutical industry: The Open PHACTS project. Antony J. Williams, williamsa@rsc.org.

3:50 – 61 – Enabling the translational medicine and drug discovery information workflow. David Evans, david.evans@reedelsevier.ch, Timothy Hoctor, Jacqui Mason, Pieder Caduff.

4:20 – 62 – Standardized drug and pharmacological class network in ontology representation. Qian Zhu, zhu.qian@mayo.edu, Cui Tao, Christopher Chute.

4:50 – 63 – Application of text mining and semantic technology in external intelligence surveillance. Yiqun H Li, lihy@lilly.com.

5:20 – Concluding Remarks

Monday, September 9, 2013

Sci–Mix – EVE Session

Indiana Convention Center
Room: Halls F&G

Jeremy Garritano, Organizer

8:00 pm – 10:00 pm

9 – ChemDraw, iPads, and collaboration tools in the classroom: Results of a joint PerkinElmer and McGraw Hill pilot at the organic chemistry undergraduate level Robin Y Smith, robin.smith@perkinelmer.com, Hans C Keil, hans.keil@perkinelmer.com, Tamara Hodge.

12 – Enabling chemistry on-the-go with modern solutions. Tamsin E Mansley, tamsin.mansley@dotmatics.com, Graeme E Dennis, Shikha O’Brien.

15 – Improving access to data: A distributed approach. Graeme E Dennis, graeme.dennis@dotmatics.com, Tamsin E Mansley, Shikha O’Brien.

22 – OPSIN: Taming the jungle of IUPAC chemical nomenclature. Daniel M Lowe, daniel@nextmovesoftware.com, Peter Murray-Rust, Robert C Glen.

26 – Molecular scaffolds are special and useful guides to discovery. Jeremy J Yang, jyang@salud.unm.edu, Cristian G Bologa, David J Wild, Tudor I Oprea.

29 – Enhanced ranking of PknB inhibitors using data fusion methods. Abhik Seal, abseal@indiana.edu, David John Wild.

30 – 3D–QSAR using quantum-mechanics-based molecular interaction fields. Ahmed El Kerdawy, ahmed.elkerdawy@chemie.uni-erlangen.de,
Stefan Güssregen, Hans Matter, Matthias Hennemann, Timothy Clark.

62 – Standardized drug and pharmacological class network in ontology representation. Qian Zhu, zhu.qian@mayo.edu, Cui Tao, Christopher chute.

64 – Teach-Discover-Treat: Round 2 competitions. Rommie E Amaro, ramaro@ucsd.edu, Johanna Jansen, Jane Tseng, Wendy Cornell.

66 – Adventures in drug discovery: For now we see through a glass, darkly. Robert C Glen, rcg28@cam.ac.uk.

83 – Teaching information literacy through an undergraduate laboratory project. Martin A Walker, walkerma@potsdam.edu.

85 – Designing instruction activities to guide students through the research lifecycle: A science librarian approach Ye Li, liye@umich.edu.

87 – Anything BUT overlooked: Librarians teaching scientific communication skills at the University of Florida. Donna T. Wrublevski, dtwrublevski@library.caltech.edu, Amy Buhler, Sara Gonzalez, Margeaux Johnson.

92 – HPCC: A suitable solution for performing drug repositioning and preclinical pharmacological profiling Arnaud Sinan Karaboga, karaboga@harmonicpharma.com, Florent Petronin, Michel Souchet, Bernard Maigret.


96 WITHDRAWN

112 – Extraction, analysis, atom-mapping, classification, and naming of reactions from pharmaceutical ELNs. Roger Sayle, roger@nextmovesoftware.com, Daniel Lowe, Noel O’Boyle.

117 – Chemotype approach to mapping the chemical landscape and exploring chemical–biological interactions within EPA’s ToxCast project. Rachelle J Bienstock, bienstock.rachelle@epa.gov, Chihae Yang, Jim Rathman, Ann M Richard.

Tuesday, September 10, 2013

Herman Skolnik Award Symposium – AM Session

Indiana Convention Center
Room: 140

Richard Cramer, Organizer
Brian Masek, Presiding

8:30 am – 11:50 am

8:30 – Introductory Remarks

8:35 – 65 – Adventures in drug discovery: For now we see through a glass, darkly. Robert C Glen, rcg28@cam.ac.uk.

9:05 – 66 – Adventures in drug discovery: For now we see through a glass, darkly. Robert C Glen, rcg28@cam.ac.uk.

9:35 – 67 – Three paradigm shifts in computer-assisted drug design: The inventors and by-standers. Yvonne C Martin, yvonnemartin@comcast.net.

10:05 – Intermission

10:50 – 69 – Evolution of QSAR from regression analysis to physical modeling. Ajay N Jain, ajain@jainlab.org.


Tuesday, September 10, 2013

Herman Skolnik Award Symposium – PM Session

Indiana Convention Center
Room: 140

Richard Cramer, Organizer
Terry Stouch, Presiding

2:00 pm – 4:45 pm

2:00 – Introductory Remarks


2:35 – 72 – Think local, act global: Some challenges in cheminformatics and drug research. Tudor I Oprea, toprea@salud.unm.edu.

3:05 – Intermission

3:20 – 73 – From library design to off-target prediction: A wide array of topomer applications. Bernd Wendt, bernd.wendt@certara.com.


4:35 – Award Presentation

Wednesday, September 11, 2013

Exchangeable Molecular and Analytical Data Formats and their Importance in Facilitating Data Exchange – AM Session Exchanging Molecular Data

Indiana Convention Center
Room: 140

Cosponsored by COMP

Antony Williams, Robert Lancashire, Organizers
Antony Williams, Presiding

8:05 am – 12:00 pm

8:05 – Introductory Remarks

8:10 – 75 – Cheminformatics runs on molfiles and its siblings: There is a molfile for that. Keith T Taylor, keith.taylor@accelrys.com.

8:40 – 76 – Reading and writing molecular file formats for data exchange of small molecules, biopolymers, and reactions Roger A Sayle, roger@nextmovesoftware.com.

9:10 – Intermission

9:20 – 77 – Facilitating accurate chemical data interconversion using Open Babel: The good, the bad, and the painful. Geoffrey R Hutchison, geoffh@pitt.edu.

9:50 – 78 – Exchanging chemical structures and ELN data: CDX, CDXML, and other formats. David Gosalvez, Alex Jewett, Phil McHale, phil.mchale@perkinelmer.com, Churl Oh, Rudy.
Chemical Information Bulletin, 2013, 65 (3) 48

Potenzone, rudy.potenzone@perkinelmer.com, Chris Strassel.

10:20 – 79 – InChI: Recent developments in the worldwide chemical structure identifier standard. Stephen Heller, steve@hellers.com.

10:50 – Intermission

11:00 – 80 – Data exchange caveats and particulars, the devil is in the details. Evan Bolton, bolton@ncbi.nlm.nih.gov.

11:30 – 81 – Pistoia Alliance and the emerging HELM standard at the “dawn of the ADC informatics era”. Sergio Rotsein, John Wise, Claire Ballamy, Michael Braxenthaler, Barry Bunin, bbunin@collaborativedrug.com.

Wednesday, September 11, 2013

Before and After Lab: Instructing Students in 'Non–Chemical' Research Skills – AM Session

Indiana Convention Center
Room: 141

Cosponsored by CHED, YCC
Andrea Twiss–Brooks, Charles Huber, Organizers
Charles Huber, Presiding

8:10 am – 12:00 pm

8:10 – Introductory Remarks

8:15 – 82 – Chemical information across San Diego County: A community college and university library collaboration for an independent synthesis project. Teri M Vogel, tmvogel@ucsd.edu, Cynthia B Gilley.

8:45 – 83 – Teaching information literacy through an undergraduate laboratory project. Martin A Walker, walkermaneumich.edu.

9:15 – 84 – Integrating citations as a teaching element into chemistry information literacy training sessions. Shu Guo, guo1s@cmich.edu.

9:45 – Intermission

10:00 – 85 – Designing instruction activities to guide students through the research lifecycle: A science librarian approach. Ye Li, liye@umich.edu.

10:30 – 86 – "I can just copy this, right?: Introducing students to copyright. Charles F. Huber, huber@library.ucsb.edu.

11:00 – 87 – Anything BUT overlooked: Librarians teaching scientific communication skills at the University of Florida. Donna T. Wrublewski, dtwrublewski@library.caltech.edu, Amy Buhler, Sara Gonzalez, Margeaux Johnson.

11:30 – 88 – Introducing Electronic Laboratory Notebooks (ELNs) to students and researchers at the University of Maryland College Park
Svetla Baykoucheva, sbykouc@umd.edu, Lee Friedman.

Wednesday, September 11, 2013

Computational Profiling and Repositioning as Promising New Ways of Drug Development – AM Session

Indiana Convention Center
Room: 142

Cosponsored by COMP
Andrew Hopkins, Violeta Isabel Perez Nueno, Organizers
Violeta Isabel Perez Nueno, Presiding
8:00 am – 12:15 pm

8:00 – Introductory Remarks

8:05 – 89 – Actual and predicted target and activity profiles for pharmaceuticals. John P Overington, jpo@ebi.ac.uk.

8:30 – 90 – Ligand promiscuity or protein redundancy?: Lessons from the PDB Esther Kellenberger, ekellen@unistra.fr, Noé Sturm, Jérémy Desaphy, Didier Rognan.

8:55 – 91 – Predicting and testing drug off- and on-targets. Brian Shoichet, bshoichet@gmail.com.

9:20 – Intermission


10:45 – Intermission


11:20 – 96 – Drug repositioning for discovery of novel TRAF2 and NCK–interacting kinase (TNIK) inhibitors. Lu Chen, LChen8@mdanderson.org, Shuxing Zhang.

11:45 – 97 – Structure–based discovery of prescription drugs that interact with Solute Carrier (SLC) transporters. Avner Schlessinger, avner.schlessinger@mssm.edu.

12:10 – Concluding Remarks

Wednesday, September 11, 2013

Exchangeable Molecular and Analytical Data Formats and their Importance in Facilitating Data Exchange – PM Session Exchanging Analytical Data

Indiana Convention Center
Room: 140

Cosponsored by COMP

Antony Williams, Robert Lancashire, Organizers
Robert Lancashire, Presiding

1:30 pm – 4:15 pm

1:30 – 98 – 30 years of JCAMP–DX formats and still going strong. Antony N Davies, Robert J Lancashire, robert.lancashire@uwimona.edu.jm.

2:00 – 99 – Knowledge sharing or what I learned in first grade. Michael Boruta.
2:30 – 100 – Long wait for exchangeable data formats vs. the evolution of data. 
Clemens Anklin, Clemens.Anklin@bruker-biospin.com.

3:00 – Intermission

3:10 – 101 – Leveraging the AnIML specification for analytical data exchange. 
Stuart J Chalk, schalk@unf.edu.

3:40 – 102 – JCAMP–MOL: A JCAMP–DX extension to allow interactive model/spectrum exploration using Jmol and JSpecView. Robert M Hanson, hansonr@stolaf.edu, Robert J Lancashire.

4:10 – Concluding Remarks

Wednesday, September 11, 2013

Before and After Lab: Instructing Students in 'Non–Chemical' Research Skills – PM Session

Indiana Convention Center Room: 141

Cosponsored by CHED, YCC
Andrea Twiss–Brooks, Charles Huber, Organizers
Charles Huber, Presiding

1:00 pm – 3:05 pm

1:00 – Introductory Remarks

1:05 – 103 – Teaching chemical information in bulk: Incorporating information skills in a large laboratory class 
Judith N. Curran, curran@pobox.upenn.edu.

1:35 – 104 – Social profile of a chemist online: The potential profits of participation. 
Antony J Williams, williamsa@rsc.org.

2:05 – 105 – Safety outreach to the academic chemistry community. 
Ralph Stuart, rstuart@cornell.edu.

2:35 – 106 – Other skills for post-graduates. Pamela J Scott, pamela.j.scott@pfizer.com.

Back to the Future: Print Resources in a Digital World – PM Session

Indian Convention Center Room: 141

Grace Baysinger, Organizer
Grace Baysinger, Presiding

3:10 pm – 5:30 pm

3:10 – Introductory Remarks

3:15 – 107 – Digitizing documents to provide a public spectroscopy database. 
Antony J Williams, williamsa@rsc.org, Colin Batchelor, William Brouwer, Valery Tkachenko.

3:45 – 108 – Whither the books? Managing access to print resources in an academic elibrary. Leah R McEwen, lr1@cornell.edu.

4:15 – Intermission

4:30 – 109 – Keeping the books on campus: The University of Chicago approach to library collection. Andrea Twiss-Brooks, atbrooks@uchicago.edu.

5:00 – 110 – Challenges and opportunities for academic research chemistry
collections in the 21st century. Grace Baysinger, graceb@stanford.edu.

Thursday, September 12, 2013

Exchangeable Molecular and Analytical Data Formats and their Importance in Facilitating Data Exchange – AM Session Data Standards to Support Publishing and Lab Notebooks

Indiana Convention Center
Room: 140

Cosponsored by COMP

Antony Williams, Robert Lancashire, Organizers
Antony Williams, Presiding

9:05 am – 10:40 am

9:05 – Introductory Remarks

9:10 – 111 – Importance of standards for data exchange and interchange on the Royal Society of Chemistry eScience platforms. Antony J. Williams, williamsa@rsc.org, Colin Batchelor, Jon Steele, Valery Tkachenko.

9:40 – 112 – Extraction, analysis, atom–mapping, classification, and naming of reactions from pharmaceutical ELNs. Roger Sayle, roger@nextmovesoftware.com, Daniel Lowe, Noel O’Boyle.

10:10 – 113 – How standards helped RSC to create The Merck Index Online. Richard Kidd, kiddr@rsc.org.

Thursday, September 12, 2013

Exchangeable Molecular and Analytical Data Formats and their Importance in Facilitating Data Exchange – AM Session Data Exchange Standards in Chemistry and Drug Discovery

Indiana Convention Center
Room: 140

Cosponsored by COMP

Antony Williams, Robert Lancashire, Organizers
Antony Williams, Presiding

10:55 am – 12:00 pm

10:55 – 114 – Practical open data exchange formats for open organic chemistry projects. Jean-Claude Bradley, bradlejc@drexel.edu, Andrew SID Lang, Antony J Williams.

11:25 – 115 – Semantic mining and prediction for drug discovery. Bin Chen, Bing He, Ying Ding, dingying@indiana.edu, David Wild.

11:55 – Concluding Remarks

Thursday, September 12, 2013

General Papers – PM Session

Indiana Convention Center
Room: 140

Jeremy Garritano, Organizer
Jeremy Garritano, Presiding

1:00 pm – 2:30 pm

1:00 – 116 – Eureka Research Workbench: A semantic approach to an open source electronic laboratory notebook. Stuart J Chalk, schalk@unf.edu.

1:30 – 117 – Chemotype approach to mapping the chemical landscape and
exploring chemical-biological interactions within EPA’s ToxCast project. Rachelle J Bienstock, bienstock.rachelle@epa.gov, Chihae Yang, Jim Rathman, Ann M Richard.

2:00 pm 118 WITHDRAWN

COSPONSORED SYMPOSIA

CHAS, Linda Stroud, Debbie Decker

Sunday, September 8, 2013

New Horizons in Chemical Health and Safety – PM Session

Indiana Convention Center
Room: 115

Cosponsored by CCS, CINF

Ralph Stuart, Organizer
Ralph Stuart, Presiding

2:50 pm – 6:40 pm

2:50 – Interductory Remarks

2:55 – 4 – Change is the only constant. Kenneth P. Fivizzani.

3:15 – 5 – How the CPT can help safety take its rightful place in the chemistry curriculum. David C Finster.

3:35 – 6 – Shifting from specific subject-matter experts to all-hazards’ safety outreach staff in EHS. Erik A Talley.

3:55 – Intermission

4:05 – 7 – New horizons, lessons from the Nanotechnology Research Center: Understanding and managing laboratory health and safety of a new or emerging technology. Laura L Hodson, Charles Geraci.


5:05 – Intermission


5:35 – 11 – Accessing chemical health and safety data online using Royal Society of Chemistry resources. Antony J. Williams, Colin Batchelor, Alexey Pshenichnov, Valery Tkachenko.


6:35 – Concluding Remarks

Cosponsored Symposia COMP, Emilio Esposito, Scott Wildman Sunday,

September 8, 2013

20 Years of PME: Symposium in Honor of Tom Darden, Lee G. Pedersen, and Darrin York – AM Session
Gerardo Cisneros, Jean-Philip Piquemal, Celeste Sagui, Organizers
Gerardo Cisneros, Tamar Schlick, Presiding

8:30 am – 11:45 am

8:30 – Introductory Remarks

8:35 – 13 – Modeling halogen compounds with a polarizable multipole force field. Pengyu Ren, Xiaojia Mu, Qiantao Wang, Lee-Ping Wang.


9:35 – 15 – Tales from a member of the PME generation: From QM/MM to GPUs. Ross C Walker.

10:05 – Intermission


10:45 – 17 – Simulations of nucleic acids and/or how PME changed my life. Thomas E. Cheatham.

11:15 – 18 – PME on the Big Iron, T3D to GPU and beyond. Michael F. Crowley.

COMP, Emilio Esposito, Scott Wildman

Sunday, September 8, 2013

20 Years of PME: Symposium in Honor of Tom Darden, Lee G. Pedersen, and

Darrin York – PM Session

Indiana Convention Center
Room: 137

Cosponsored by BIOL, CINF, MEDI, PHYS

Celeste Sagui, Gerardo Cisneros, Jean-Philip Piquemal, Organizers
Pengyu Ren, Ross Walker, Presiding

1:30 pm – 4:45 pm

1:30 – 47 – Protein structure refinement with a polarizable force field and PME. Stephen LuCore, Jooyeon Park, Kyle Powers, Shuxiang Li, Timothy D Fenn, Michael J Schnieders.

2:00 – 48 – Atomistic dynamics simulations of DNA polymerase conformational pathways help unravel mechanistic aspects of enzyme repair. Yunlang Li, Tamar Schlick.

2:30 – 49 – All-atom simulations of biomolecular crystals. David A Case.

3:00 – Intermission


COMP, Emilio Esposito, Scott Wildman

Monday, September 9, 2013

20 Years of PME: Symposium in Honor
of Tom Darden, Lee G. Pedersen, and Darrin York – AM Session

Indiana Convention Center
Room: 137

Cosponsored by BIOL, CINF, MEDI, PHYS

Celeste Sagui, Gerardo Cisneros, Jean-Philip Piquemal, Organizers
David Case, Jean-Philip Piquemal, Presiding

8:30 am – 11:45 am

8:30 – 78 – On the numerical accuracy of Ewald, smooth particle mesh Ewald, and staggered mesh Ewald methods for inhomogeneous and correlated molecular systems. Han Wang, Pingwen Zhang, Christof Schütte.

9:00 – 79 – Particle-mesh Ewald summation for long-range electrostatics in biomolecular simulations: How important are long-range electrostatics corrections? Terry P Lybrand.


10:00 – Intermission


10:45 – 82 – Linear scaling density functional tight binding models. Timothy J Giese, Darrin M York, Erich Kuechler.

11:15 – 83 – Systematic improvement on the classical molecular model of water

Lee-Ping Wang, Teresa Head-Gordon, Jay W. Ponder, Pengyu Ren, John D. Chodera, Peter K. Eastman, Todd J. Martinez, Vijay S. Pande.

COMP, Emilio Esposito, Scott Wildman

Monday, September 9, 2013

20 Years of PME: Symposium in Honor of Tom Darden, Lee G. Pedersen, and Darrin York – PM Session

Indiana Convention Center
Room: 137

Cosponsored by BIOL, CINF, MEDI, PHYS

Celeste Sagui, Gerardo Cisneros, Jean-Philip Piquemal, Organizers
Celeste Sagui, Michael Schnieders, Presiding

1:30 pm – 4:45 pm


2:30 – 112 – Non-equilibrium free energy methods for biomolecular simulations. Celeste Sagui, Volodymyr Babin, Mahmoud Moradi, Christopher Roland.

3:00 – Intermission


3:45 – 114 – Periodic boundary conditions
in molecular simulations. Gerhard Hummer.


COMP, Emilio Esposito, Scott Wildman

Tuesday, September 10, 2013

Developments in Systems Chemical Biology – PM Session

Indiana Convention Center
Room: 136

Cosponsored by CINF, MEDI
Jeffrey Sutherland, Organizer
Jeffrey Sutherland, Michal Vieth, Presiding

1:30 pm – 5:45 pm


3:00 – 169 – Network-focused chemogenomics in phenotypic HTS. Jeremy L Jenkins.

3:30 – Intermission


COMP, Emilio Esposito, Scott Wildman

Monday, September 9, 2013

Drug Discovery – PM Session
We Seek Information to Solve our #structuralBioinformatics and #target–basedDesign Ponderings #drugDisco

Indiana Convention Center
Room: 136

Cosponsored by BIOL, CINF, MEDI
Scott Wildman, Y. Jane Tseng, Organizers
Y. Jane Tseng, Presiding

1:30 pm – 5:15 pm


2:00 – 104 – Impact of inhibitors on the tertiary structure of acetylcholinesterase. Emilio Xavier Esposito, Terry R Stouch, Troy Wymore, Jeffry D Madura.

2:30 – 105 – Structure-based function
prediction for proteins of unknown function employing ProBiS enhanced by molecular dynamics. Dusanka Janezic

3:00 – 106 – Molecular dynamics studies of the role of protein flexibility in immunological molecular recognition Cory M Ayres, Daniel R Scott, Brian M Baker, Steven A Corcelli.

3:30 – Intermission


COMP, Emilio Esposito, Scott Wildman

Tuesday, September 10, 2013
Drug Discovery – AM Session It’s What You do with the Information... #StructuralBioinformatics #Target-basedDesign #drugDisco #

Indiana Convention Center
Room: 136

Cosponsored by BIOL, CINF, MEDI

Scott Wildman, Y. Jane Tseng, Organizers
Terry Stouch, Presiding

8:30 am – 11:45 am

8:30 – 136 – Identification of allosteric Akt inhibitors by computational methods.

Oya Gursoy, E1f Oezkirimi, Kutlu O Ulgen.

9:00 – 137 – Characterizing an evolved CXC chemokine receptor interface. Sara E Nichols, Carlos X Hernández, Yi Wang, J. Andrew McCammon.


10:00 – Intermission


11:15 – 141 – Automation of Monte Carlo relative free energy calculations for ligand design. Julien Michel, Hannes Loeffler, Christopher Woods.

COMP, Emilio Esposito, Scott Wildman

Wednesday, September 11, 2013
Drug Discovery – AM Session It’s not @acsCOMP without a Little #methodMadness! #drugDisco Style

Indiana Convention Center
Room: 136

Cosponsored by BIOL, CINF, MEDI

Scott Wildman, Y. Jane Tseng, Organizers
Ryan Coleman, Presiding
8:30 am – 11:45 am


9:00 – 277 – Comprehensive validation of a computational docking methodology to identify the noncovalent binding site of a ligand to DNA. Christos Deligkaris, Anthony T. Ascone, Kevin J. Sweeney.


10:00 – Intermission


Room: 136

Cosponsored by BIOL, CINF, MEDI

Scott Wildman, Y. Jane Tseng, Organizers

Magdalena Korczynska, Presiding

8:30 am – 11:45 am

8:30 – 342 – WITHDRAWN


10:00 – Intermission


10:45 – 346 – Finding and using activity cliffs in 3D: Gaining more SAR information during lead optimization. Tim Cheeseright, Rae Lawrence, Mark Mackey, Martin Slater.


COMP, Emilio Esposito, Scott Wildman

Thursday, September 12, 2013

Drug Discovery – AM Session

Analyzing the Chemical and Structural Information for #drugDisco Keeps You Ahead of the Rest. Not Really Sure who the Rest are...

Indiana Convention Center
Sunday, September 8, 2013

Emerging Technologies in Computational Chemistry – PM Session

Indiana Convention Center
Room: 136

Cosponsored by BIOL, CINF, MEDI, PHYS

Curtis Breneman, Organizer
Curtis Breneman, Presiding

1:30 pm – 4:55 pm

1:30 – Introductory Remarks

1:35 – 40 – Objective models for ligand binding characteristics from quantum mechanics. Timothy Clark.


2:50 – Intermission


3:30 – 44 – Catching transition states by throwing ropes to stable intermediates. Paul Zimmerman.


4:45 – Award Presentation

COMP, Emilio Esposito, Scott Wildman

Tuesday, September 10, 2013

Expanding Computational Chemistry with GPUs – AM Session

Indiana Convention Center
Room: 137

Cosponsored by BIOL, CINF, MEDI, PHYS

Kenneth Merz, Scott Le Grand, Organizers

8:30 am – 11:20 am

8:30 – Introductory Remarks

8:35 am 142 WITHDRAWN


10:05 – Intermission

10:20 – 145 – Quantum chemistry many-body methods on multicore processors and accelerators. Eugene DePrince, Jeff.
Hammond.

10:50 – 146 – Interactive ab initio molecular dynamics. Nathan Luehr, Alex Jin, Todd J Martinez.

COMP, Emilio Esposito, Scott Wildman

Tuesday, September 10, 2013

Expanding Computational Chemistry with GPUs – PM Session

Indiana Convention Center
Room: 137

Cosponsored by BIOL, CINF, MEDI, PHYS

Kenneth Merz, Scott Le Grand, Organizers
Kenneth Merz, Scott Le Grand, Presiding

1:30 pm – 4:20 pm

1:30 – Introductory Remarks

1:35 – 174 – Apples vs. oranges or apples vs. apples? A comparison of the real–time, real–space electronic structure approach vs. the Gaussian basis set method using general purpose graphics processing units (GPU). Alan Aspuru–Guzik.

2:05 – 175 – GPU-enabled molecular dynamics on desktop: From the classical to reactive force field. Xiaoxia Li, Mo Zheng, Jian Liu, Zhongliang Liu, Xiaofang Tao, Li Guo.

2:35 – 176 – Acceleration of electron repulsion integral evaluation on graphics processing units via use of recurrence relations. Yipu Miao, Kenneth M Merz.

3:05 – Intermission


COMP, Emilio Esposito, Scott Wildman

Wednesday, September 11, 2013

Expanding Computational Chemistry with GPUs – AM Session

Indiana Convention Center
Room: 137

Cosponsored by BIOL, CINF, MEDI, PHYS

Kenneth Merz, Scott Le Grand, Organizers
Kenneth Merz, Scott Le Grand, Presiding

8:30 am – 11:45 am

8:30 – 282 – Unlikely symbiosis: How supercomputing and gaming are evolving together. Sarah Tariq.

9:00 – 283 – Biological sequence alignment with GPU-BLAST. Tianluo Chen, Nick Sahinidis, Panos Vouzis.


10:00 – Intermission

10:15 – 285 – Simulating millisecond dynamics in all-atom detail using a GPU cluster: Protein folding and conformational
change of kinases and GPCRs. Vijay Pande.


COMP, Emilio Esposito, Scott Wildman

Sunday, September 8, 2013

Material Science – AM Session Loving the #matSci research? We are. Follow us on twitter! @acsCOMP

Indiana Convention Center Room: 136

Cosponsored by CINF, PHYS, PMSE, POLY

Maciej Haranczyk, Organizer
Luke Achenie, Presiding

8:30 am – 11:45 am


9:00 – 8 – Effect of catalysts and solvent on the stereoregularity and kinetics of acrylate derivatives: A computational approach. Viktorya Aviyente, Tugba Furuncuoglu Ozaltin.


10:00 – Intermission


COMP, Emilio Esposito, Scott Wildman

Monday, September 9, 2013

Material Science – AM Session The Great Thing about #matSci is its Use of Different Computational Methods #methodMadness

Indiana Convention Center Room: 136

Cosponsored by CINF, PHYS, PMSE, POLY

Maciej Haranczyk, Organizer
Rebecca Weber, Presiding

8:30 am – 11:45 am

8:30 – 72 – Quantum chemical molecular dynamics study on film growth mechanisms of microcrystalline silicon solar cells. Takuya Kuwahara, Hiroshi Ito, Yuji Higuchi, Nobuki Ozawa, Momoji Kubo.
9:00 – 73 – Combined experimental and theoretical approach to identify crystal structures of new high-pressure polymorphs of FOX–7 crystal. Boris Averkiev, Yuchuan Tao, Santanu Chaudhuri, Zbigniew A Dreger, Yogendra M Gupta.


10:00 – Intermission


COMP, Emilio Esposito, Scott Wildman

Wednesday, September 11, 2013

Membranes – PM Session @ acsCOMPprog Loves a Good Pun but You Won’t Find One Here. Just Membranes. All Afternoon #membranes

Indiana Convention Center
Room: 137
Cosponsored by BIOL, CINF, PHYS

Michael Feig, Organizer

Vahid Mirjalili, Presiding

1:30 pm – 5:15 pm


2:00 – 313 – Steered molecular dynamics provide insights into the translocation of the cell penetrating peptide pVEC Begum Alaybeyoglu, Elif Ozkirimli.


3:00 pm 315 WITHDRAWN

3:30 – Intermission


4:15 pm 317 WITHDRAWN


COMP, Emilio Esposito, Scott Wildman

Sunday, September 8, 2013

Molecular Mechanics – AM Session Kicking off #ACSindy with #molMech Applications. Remember to Follow @ acsCOMP and @acsCOMPprog
8:30 am – 11:45 am


9:00 – 27 – Efficient method to elucidate ligand binding pathways: Binding kinetics studies for cryptophane host-guest complexes. **Chia-en A. Chang**, Christopher C Roberts.


10:00 – Intermission


COMP, Emilio Esposito, Scott Wildman

Monday, September 9, 2013

Molecular Mechanics – AM Session
Electrosstatics and Polarization Help
Things Stick Together. Hopefully You'll
Stick Around all Morning #molMech

COMP, Emilio Esposito, Scott Wildman

Tuesday, September 10, 2013

Molecular Mechanics – AM Session
Exploring Proteins with #molMech is Part of a Heart Healthy Morning. Not Really. But your Mind will Thank you. #expandYourBrain

Indiana Convention Center
Room: 139

Cosponsored by BIOL, CINF, MEDI, PHYS

Michael Feig, Organizer
Eduardo Sproviero, Presiding

8:30 am – 11:45 am


9:00 – 155 – Mechanisms of small molecule binding to intrinsically disordered proteins. Julien Michel, Rémi Cuchillo.


10:00 – Intermission


11:15 – 159 – Molecular mechanism of how palmitate affects the activity of IRE1α/ PKR. Liang Fang, Hyunju Cho, Christina Chan, Michael Feig.

COMP, Emilio Esposito, Scott Wildman

Wednesday, September 11, 2013

Molecular Mechanics – AM Session
Come See if Your Favorite Protein is Part of this #molMech Session!

Indiana Convention Center
Room: 139

Cosponsored by BIOL, CINF, MEDI, PHYS

Michael Feig, Organizer
Monica Berrondo, Presiding

8:30 am – 11:45 am


9:00 – 295 – Integrative modeling of complex biological assemblies in transcription. Ivaylo Ivanov.

9:30 – 296 – Novel PIP2 lipid binding to focal adhesion kinase probed through coarse-grained molecular dynamics simulations. Jun Feng, Blake Mertz.

10:00 – Intermission


Sunday, September 8, 2013

Current Challenges in Cheminformatics: Exploiting Information and Knowledge in Structured and Unstructured Environments – AM Session

Cheminformatics Tools and Methods

Indiana Convention Center
Room: 140
Cosponsored by COMP

Neil Kirby, Dirk Tomandl, Organizers
Dirk Tomandl, Presiding

8:10 am – 12:00 pm

8:10 – Introductory Remarks

8:15 – 1 – New insights and tools for capturing, validating, and utilizing structure/property data: Curiously similar approaches for drug discovery and materials science. Curt M Breneman, brenec@rpi.edu, Ke Wu, Lisa Morkowchuk, Jed Zaretzki, Sourav Das, Michael Krein.

Department of Chemistry & Chemical Biology, Rensselaer Polytechnic Institute, Troy, NY 12180, United States

As with all “data driven” areas of scientific inquiry, access to reliable, accurate and well-curated experimental data (accompanied by appropriate meta-data) is crucial for exploiting the information embedded in the data to create valid, useful predictive models. In reality, access to such large, high-quality data sets can be difficult for practical reasons. In this talk, examples of model building and validation using public domain and/or IP-protected data will be discussed for several problem domains, including the prediction of CYP sites of metabolism, off-target drug interaction predictions, and materials informatics applications. Emphasis will be placed on both the commonalities of the problems, as well as their differences.

Lilly Research Laboratories, Eli Lilly and Company, Indianapolis, Indiana (IN) 46285, United States

A main objective of quantitative structure–activity relationship (QSAR) is to elucidate the contribution of structural features to an observed biological property and, thus, enable the application of such knowledge in subsequent drug discovery efforts. Essentially, QSAR is a screening data–driven method aiming to extract knowledge for biological data interpretation and, for predictive modeling purposes. In a typical scenario, QSAR models are developed on an assay screening dataset of a specific target of interest and are then used for understanding chemical structure–activity relations or, as filters, i.e. to predict the biological behavior of unknown molecules on the same or a related target. The overwhelming majority of QSAR models have a limited life span. Some exceptions do exist, with ‘global’ QSAR models that are repeatedly used to predict certain properties, often related to ADMET. Research efforts in the QSAR field have largely focused on the investigation of predictive modeling algorithms, molecular descriptors, similarity measures and training data collection and pre-processing. More recently, efforts in assessing and expanding the applicability domain of models as well as applications for selectivity prediction have been reported in the literature. A potentially useful yet unexplored research direction is that of SAR summarization and the identification of ‘persistent’ SAR transferable across screening datasets and, potentially, targets of interest. In this presentation we present our initial results in summarizing SAR knowledge from screening datasets and identifying specific SAR with ‘cross–assay’ potential. We introduce a method for the detailed encoding and storage of SAR using scaffold-based analysis and reasoning. Techniques for SAR knowledge comparison and applicability analysis for potential reuse are also described. Results from the application of the approach to multiple datasets are presented as are potential uses of the method. A discussion on lessons learned, issues to be resolved, and future development directions concludes the presentation.


Scilligence Corporation, Burlington, MA 01803, United States

With prevalence of cross-organization collaborations, R&D reorganizations and company merger &acquisition, knowledge can be easily lost in a large pile of unstructured data such as PPT, Word, Excel and PDF. Through a case study example, knowledge mining and preservation by structure searching makes possible by Scilligence’s informatics tools.

9:45 – Intermission

9:55 – 4 – Computational approach of thermodynamic fragmentation applied to formulating badly soluble actives. Johannes Fraaije (1,2), j.fraaije@chem.leidenuniv.nl, Monica Bulacu (2), Ruben Gracia (2), Shyamal K Nath (2).

(1) Soft Matter Chemistry, University of Leiden, Leiden, The Netherlands, (2) Culgi, Leiden, The Netherlands

We address one of the major computational problems in the
pharmaceutical and agrochemical business: the rational design of delivery vehicles for badly soluble actives. For example, in a typical scenario in a discovery/development cycle, one has only a few months, or even weeks, to find a proper formulation, driven by very costly time expenditure way down the pipeline. The chemical space for formulations is enormous, and even if one considers modern high-throughput experimentation, chances are the optimal formulation cannot be found by experimentation alone. For chemical informatics, such design also poses quite some challenges, since classical descriptor technology is almost always unitary (component specific), and therefore not suited for the non-linear binary interactions between active and delivery matrix. The more so, since in many cases one seeks or tests delivery systems that are based on some self-assembly structure, such as an emulsion, micelle or liposome. In contrast, in existing chemical informatics methods for the much simpler logP predictions one can rely on such unitary descriptors quite well, since the (iso-octanol/water) matrix in this case is known and constant over a known training set. Our strategy is completely different. We develop computational screening technologies that do take into account the structural resolution of both active and matrix, by employing a novel concept of thermodynamic fragmentation. Thereby the interactions of drug-matrix is described as a interaction between fragments, with the advantage that the fragmentation can be calibrated by (non-pharma) engineering databases. The presentation discusses in depth the motivation, the algorithm and some highlights from recent applications.

10:25 – 5 – Spaghetti seas and filter forests: Navigating Pipeline Pilot protocol creation with maintenance in mind.

Jennifer L Heymont, Jennifer_Heymont@eisai.com.

Department of Information Technology, Eisai Inc., Andover, MA 01810, United States

Pipeline Pilot provides a unique environment that is part workflow diagram, part Lego, part plain old code and which supports extremely fast application development for informatics. The ability to code in this environment is very friendly to the beginner while also supporting complex development by experts, however the non-linear nature of the resulting protocols can lead to a maintenance nightmare, especially if more than one person is responsible for supporting a particular application. Our group has developed a set of general methods and guidelines for developing applications that can be easily maintained by a group with diverse backgrounds; these will be presented within the context of the Pipeline Pilot application.


Philip J Skinner, Philip.skinner@perkinelmer.com, Rudy Potenzzone, Megean Schoenberg, Amy Kallmerten, Phil McHale, Michael Swartz.

PerkinElmer, United States

The early development of ELNs focused on recording chemical syntheses and providing forms to structure the collection and calculation of reaction-centered data such as stoichiometry tables. Within early adopter organizations these databases now contain a trove of in-house reaction methodology, including valuable data on not only successful and but also failed or poorly working reactions, in contrast with
traditional journals. Searching ELN-based reactions with advanced analytics tools such as the TIBCO Spotfire® software, allows for the creation of structured, tabular views into the historical synthesis data. Such views were unavailable through traditional ELN search tools which generally provided only links to the experiments that matched search criteria. With Spotfire, chemists can ask complex, dynamic and structured questions of the methodology data, and visualize that data in various graphical forms. This will be exemplified through case-studies such as reaction optimization, project management and trend analysis.

11:25 – 7 – Plexus: A clean web application for structural data analysis

David Deng, ddeng@chemaxon.com, Andras Stracz.

ChemAxon LLC, Cambridge, MA 02138, United States

Plexus is a new web application developed by ChemAxon that can integrate with various discovery tools. It focuses on delivering a user-friendly interface for data visualization. It can be used to access different ChemAxon applications, including structure database management, structure search, phys-chem property calculation, virtual reaction synthesis. It will also apply to virtual library enumeration and Markush structure editor. Its simplicity is reflected by a clean web interface with Marvin for JavaScript as the structure editor, no local installation and a low learning curve. A demo site has been set up and can be access freely. In this presentation, various features will be demonstrated in compound library design, including Markush structure analysis, virtual screening and effortless reaction enumeration.

11:55 – Concluding Remarks
8:40 – 9 – ChemDraw, iPads, and collaboration tools in the classroom: Results of a joint PerkinElmer and McGraw Hill pilot at the organic chemistry undergraduate level. Robin Y Smith (1), robin.smith@perkinelmer.com, Hans C Keil (1), hans.keil@perkinelmer.com, Tamara Hodge (2).

(1) PerkinElmer, Waltham, Massachusetts 02451, United States, (2) McGraw Hill Education, Dubuque, Iowa 52001, United States

In partnership with McGraw Hill Education, PerkinElmer will conduct a pilot across several undergraduate organic chemistry classes testing the use of iPads, ChemDraw, Chem3D and a new cloud-based collaboration service. The pilot will test the effectiveness of tablet-based learning at the undergraduate level. Working closely with professors and students, PerkinElmer will adapt ChemDraw and other current software for the learning platforms and techniques of tomorrow. McGraw Hill, as a leading company in the education software industry, will work closely with participating professors to analyze the success of the pilot (versus previous classes) and make recommendations on future steps. The pilot will take place over the 2013 summer session with results being compiled in August and presented publically for the first time at the 246th ACS National Meeting.

9:10 – 10 – Can I get that to go? Reading research articles on a tablet. Jeff Lang, j_lang@acs.org. American Chemical Society Publications, Washington, D.C 20036, United States

Tablets are made for reading, but journals aren’t all made for tablets. How do we get beyond PDFs for reading journal articles on a tablet? What factors are locking readers in to their printed pages and laptops when they read a fully composed article? Learn how researchers are reacting to ACS ActiveView PDF on desktop and how that can apply to the tablet reading experience.

9:40 – Intermission

9:55 – 11 – JSmol: Full-service molecular visualization on the Web without Java. Robert Hanson (1), hansonr@stolaf.edu, Zhou Renjian (4), Takanori Nakane (2), Paul Pillot (3).

(1) Department of Chemistry, St. Olaf College, Northfield, MN 55057, United States, (2) Department of Molecular Biology, Kyoto University, Kyoto, Yoshidakonoe-cho, Sakyo-ku 606-8501, Japan, (3) l’academie d’Orleans-Tours, Orleans, France, (4) Unaffiliated, China

We have recently ported the Java-based Jmol applet to pure JavaScript, enabling full Jmol functionality on tablets and other devices that do not have Java capabilities. The implementation, JSmol (JSmol. sourceforge.net), reproduces all of the functionality of Jmol/Java, but with no Java. In this presentation we will compare performance of Jmol and JSmol on a variety of platforms, highlight their similarities and differences, and take a look at the future of web-based molecular visualization in science, education, and publishing.

10:25 – 12 – Enabling chemistry on-the-go with modern solutions. Tamsin E Mansley (1), tamsin.mansley@dotmatics.com, Graeme E Dennis (1), Shikha O’Brien (2).
The paradigm in chemical education and research is continuously evolving as technology becomes more pervasive. Even today scientists and students are dependent on a multitude of tools to capture and share data: paper notebooks, calculators, text books, desktop personal computers, etc. Are we playing catch-up to social media and technology when it comes to science? Today laptops, tablets, mobile devices and social media are commonplace and there is a need to provide technology to educators, users and early adopters through these media. This presentation will focus on our ability to support distance learning and on-the-go chemistry in research organizations through web-based and mobile applications, including free tools such as Dotmatics’ Elemental chemical sketching app for iPad, iPhone and Android devices.

10:55 – 13 – Tablets in the lab: Enabling the flow of chemical synthesis data into a chemistry repository. Simon J Coles (1), s.j.coles@soton.ac.uk, Richard J Whitby (1), Alieen Day (2), Cerys Willoughby (1), Valery Tkachenko (2), Jeremy G Frey (1), Antony J Williams (2).

(1) Chemistry, University of Southampton, Southampton, Hampshire SO17 1BJ, United Kingdom, (2) Royal Society of Chemistry, Cambridge, United Kingdom

Structures, syntheses and spectra, together with a myriad of other properties are measured in chemistry laboratories around the world - and in an escalating number. Increasingly, these data are captured and stored in electronic formats making them amenable to data sharing and searching. The Electronic Laboratory Notebook (ELN) is the new data capture of choice for many organizations, primarily utilized with the intention of securing intellectual property protection and providing improved searching and access to the data across an organisation. For several years we have been developing an ELN system, LabTrove, that not only enables this capture and curation but also provides a platform to share the information - in a selective way, as an aid to formal publication or openly on the web. We are entering a new era in terms of the willingness to share data with other scientists, generally termed “Open Data” and are only just beginning to understand the new science that this behaviour can promote. Unfortunately there is a significant bottleneck in this process - that is the physical capture of this information in the laboratory. Native capture of electronic in the synthesis lab has always been a challenge and a compromise for traditional desktop or laptop computers, however the pervasive, non cumbersome nature and simple interactivity of tablet computers has a very real potential to be adopted by chemists in the lab. Working with the Dial-a-Molecule initiative, we report on an ELN environment amenable to the capture of synthetic chemistry procedures and associated data - that is a system where a synthesis experiment can be planned on the office computer and then actions and observations recorded in the lab on a tablet. This ELN ecosystem has now also been integrated with the publicly accessible resources of the Royal Society of Chemistry (ChemSpider and ChemSpider SyntheticPages) in order to publish the data and provide access to the chemistry community.

11:25 – 14 – New strategy to engage mobile computing users and developers Steven M Muskal, smuskal@eidogensertanty.com.
Sunday, September 8, 2013

Current Challenges in Cheminformatics: Exploiting Information and Knowledge in Structured and Unstructured Environments – PM Session Data Access, Usage and Pitfalls

Indiana Convention Center
Room: 140

Cosponsored by COMP

Dirk Tomandl, Neil Kirby, Organizers
Dirk Tomandl, Presiding

1:10 pm – 3:50 pm

1:10 – Introductory Remarks

1:15 – 15 – Improving access to data: A distributed approach. Graeme E Dennis, (1), graeme.dennis@dotmatics.com, Tamsin E Mansley (1), Shikha O’Brien (2).

(1) Dotmatics, Inc., Woburn, MA 01801, United States, (2) Dotmatics, Inc., San Diego, CA 92121, United States

Today’s environment requires decentralized organizations to exchange critical information, often in a variety of formats and through multiple communication channels. Addressing issues of security, data loss and poor communication are essential for any project’s success. Among the greatest challenges faced by the scientists today are (a) dealing with the data deluge and (b) accessing and making sense of the data. How is unstructured data to be handled in a way that permits querying, browsing, and analysis while retaining all the meaning of its original presentation? Strategies permitting scientists to have secure, ready access to their own data, irrespective of where it might be located, in a format that is meaningful to them, will be presented.

1:45 – 16 – Digital fractionation: Using NoSQL technology to extract biochemical data from big data repositories. Lewis Y Geer, lewis.geer@nih.gov, Lianyi Han, Yanli Wang, Evan E Bolton, Siqian He, Bo Yu.

NCBI/NLM/NIH, Bethesda, MD 20894, United States

An exponential increase in the complexity and amount of chemical and biological information generated from large scale projects has posed challenges for traditional query engines and methods. This issue is particularly acute in databases that serve as repositories for public data, such as PubChem. Nontraditional query methods, such as NoSQL, promise to address scalability and access to tractable subsets of data while opening opportunities for discovery and analysis across disparate data sources. The PubChem DataDicer is one such query engine and we describe its architecture along with novel features and interfaces.

2:15 – Intermission

2:20 – 17 – Exploring interconnectivity of molecular biology databases and data mining tools. Stefan M Furrer, stfurrer@indiana.edu, David J Wild.

School of Informatics and Computing, Indiana University, Bloomington, IN 47408, United States

Over the past twenty years the number of
databases related to molecular biology and connected fields dramatically increased driven by science trends and enabled by computational advances. Effective navigation of highly connected databases results in extremely detailed, multifaceted and integrated datasets. Building tools that allow searches across different data sources are thus essential to mine the available information. Over the years we noticed not only a steadily increasing number of databases but also stronger data interconnectivity. We thus surveyed current existing molecular biology online databases in relation with associated data mining tools aiming to extract information and transform it into meaningful knowledge.

2:50 – 18 – Caveats in the use of pharmaceutical drug discovery data. Terry R Stouch, tstouch@gmail.com.

Science For Solutions, LLC, West Windsor, NJ - New Jersey 085505354, United States

Drug discovery data is seldom generated for use in predictice sciences and the development of global QSAR models. It is always tailored for a particular purpose. It is often highly focused in terms of range. It is best interpreted in context of closely related compounds. The temporal nature of the data can create problems in interpretation of data developed over long time periods. True error is always greater than the error of measurement. Important meta data, even if it is databased, can easily be lost during transfer or conversion of the data and is particularly a problem for capture via web-crawling. Issues important to use in predictive models will be detailed with examples from a large range of endpoints and assays and from many sources and supporting commentary by those that generated the data. Rules of thumb for error estimation will be suggested. Effects on QSAR models will be discussed.

3:20 – 19 – Extracting synthetic knowledge from reaction databases. Orr Ravitz (1), ravitz@simbiosys.com, James Law (1), Anthony Cook (2), A. Peter Johnson (2).

(1) SimBioSys Inc., Toronto, Ontario M9W 6V1, Canada, (2) School of Chemistry, University of Leeds, Leeds, United Kingdom

Underpinning the computer-aided synthesis design system, ARChem, are algorithms that extract synthetic knowledge from large reaction databases. Reaction rules that facilitate retrosynthetic analysis are generated automatically by capturing the essence of the individual reactions in the database, clustering together examples that share the same underlying chemistry, and making generalizations based on electronic properties of functional groups and moieties in the reactants and products. Each cluster of examples is also used to derive information about expected yields, regioselectivity, functional group compatibility, and stereochemistry. At a higher level, a hierarchy of rules is constructed as a curatorial and administrative tool. The automated extraction of synthetic knowledge is crucial for capturing the full breadth of the methods encapsulated in the database, but it also allows us to gain further insights on various aspects of organic chemistry as well as on the data source itself. For example, analysis of electronic patterns in electrophilic substitution reactions offers new observations regarding regioselectivity, and rule-hierarchy suggests a chemistry-driven reaction classification approach. In this talk we will describe some of the main computational approaches for knowledge extraction that
are used in ARChem, and will present results validating known principles as well as results exposing less obvious properties. The limitations of reaction databases as well as of the current algorithms will be discussed.

Sunday, September 8, 2013

Current Challenges in Cheminformatics: Exploiting Information and Knowledge in Structured and Unstructured Environments – PM Session Chemical Structures in Documents

Indiana Convention Center
Room: 140

Cosponsored by COMP

Dirk Tomandl, Neil Kirby, Organizers
Dirk Tomandl, Presiding

4:00 pm – 5:30 pm

4:00 – 20 – Validation and characterization of chemical structures derived from names and images in scientific documents. John B Kinney, john.b.kinney@dupont.com. Crop Protection, DuPont, Newark, Delaware 19714, United States

High-quality software applications currently convert chemical names and images into chemical structures on a large scale to automatically curate large document sets such as the patent corpus. The challenge remains however, in the accuracy of the assigned structures. Errors can arise from inconsistencies in the quality of the original source documents. For example, OCR and typographical errors in the text or pixelated and fuzzy lines in the images all contribute to uncertainty in assigning structures. To address this challenge we are working in collaboration with IBM and several other companies to verify the structures of the millions of unique chemical entities extracted from these documents. This talk will discuss processes that we have developed in order to characterize and validate the structures that are identified by the image and text conversion algorithms as well as progress made in linking the structures to the context of the original text.

4:30 – 21 – Tackling the difficult areas of chemical entity extraction: Misspelt chemical names and unconventional entities. Daniel M Lowe, daniel@nextmovesoftware.com, Roger A Sayle. NextMove Software Ltd, Cambridge, Cambridgeshire CB4 0EY, United Kingdom

Extracting the structures of small molecules from unstructured text is now a mature field, however there still remain areas that present considerable difficulty or have until this point remained unexplored. One such area is identification of chemical names with misspellings or errors introduced by optical character recognition. The approach we have taken employs a formal grammar describing the syntax of a systematic name. To provide coverage over the vast majority of organic nomenclature including carbohydrates, amino acids and natural products we have developed a new way of representing the grammar such as to allow an order of magnitude more states than previous efforts1 whilst simultaneously reducing memory consumption. To efficiently perform spelling correction against this grammar we will describe a heuristic spelling correction algorithm. Another area that remains underexplored is the identification and resolution of chemical line formulae by which we also include domain specific line formulae such as are used to describe oligosaccharides and
peptides. We describe the recognition and resolution of these often overlooked chemical entities. We also show how one can identify entities such as journal and patent references, which can aid in the navigation of semantically enhanced documents. (1) Sayle, R.; Xie, P. H.; Muresan, S. Improved Chemical Text Mining of Patents with Infinite Dictionaries and Automatic Spelling Correction. *J. Chem. Inf. Model.* **2011**, 52, 51–62.

5:00 – 22 – OPSIN: Taming the jungle of IUPAC chemical nomenclature

Daniel M Lowe, daniel@nextmovesoftware.com, Peter Murray-Rust, Robert C Glen.

*Department of Chemistry, University of Cambridge, Cambridge, Cambridgeshire CB2 1EW, United Kingdom*

OPSIN (Open Parser for Systematic IUPAC Nomenclature) is an open source freely available program for converting chemical names, especially those that are systematic in nature, to chemical structures. The software is available as a Java library, command-line interface and as a web service (opsin.ch.cam.ac.uk). OPSIN accepts names that conform to either IUPAC or CAS nomenclature and can convert them to SMILES, InChI and CML (Chemical Markup Language). OPSIN has grown from covering only simple general organic chemical nomenclature to the point of having competent coverage of all areas of organic chemical nomenclature. One of the most recent additions is comprehensive support for the nomenclature of carbohydrates. This brings support for dialdoses, diketoses, ketoaldoses, alditols, aldonic acids, uronic acids, aldaric acids, glycosides and oligosaccharides, in both the open chain and cyclic forms, named systematically or from trivial sugar stems with support for modification terms such as anhydro or deoxy. OPSIN’s support for specialised and general organic nomenclature will be demonstrated through illustrative examples and accompanying performance metrics. We focus in particular on areas of nomenclature for which support was recently added and those that are complex to implement such as fused ring nomenclature.

Sunday, September 8, 2013

**Graduate Student Research Symposium in Cheminformatics, Information Science, and Library Science – PM Session**

Indiana Convention Center
Room: 141

Cosponsored by COMP

Gary Wiggins, Organizers
Gary Wiggins, Presiding

2:00 pm – 5:25 pm

2:00 – Introductory Remarks

2:05 – 23 – Open chemical information: Comparison of Chemical Information Sources (CIS) Wikibook and eXplore Chemical Information Teaching Resources (XCITR). Yan He, yh4@iuuk.edu.

*Indiana University Kokomo, Kokomo, IN 46074, United States*

This presentation will compare the scope and features of two open chemical information resources: Chemical Information Sources (CIS) Wikibook and eXplore Chemical Information Teaching Resources (XCITR). A part of
the presentation will be a tutorial on how
to add pages and images or edit existing
pages in the wikibook.

2:35 – 24 – Retractions in chemistry:
Prevalence and impact. Elsa Alvaro,
ealvaro@indiana.edu.

SLIS and Chemistry Library, Indiana
University, Bloomington, Indiana 47405,
United States

Scientific research is, in most cases,
cumulative. New contributions are built
upon prior research published mainly in
the scientific literature. In order to preserve
the integrity of the scientific record,
it is important that articles containing
invalid research are retracted. In this
communication, we will study retracted
articles in the chemical literature. We
will focus on two different aspects. First,
we will examine the characteristics of
the phenomenon, including rate, time to
retraction, and reasons for retraction, and
we will establish a comparison between
chemistry and other disciplines. Then,
we will determine the impact of retracted
articles in chemistry by analyzing their
propagation to the subsequent literature.
To understand the invalid research flow we
will apply social network theory principles.

3:05 – 25 – Informatics tools for interacting
with literature and chemical databases to
build pharmacological networks of drug-
induced neuropathy. Junguk Hur (1,3),
juhur@umich.edu, Abra Guo (2), Eva L.
Feldman (1), Jane P.F. Bai (3).

(1) Department of Neurology, University
of Michigan, Ann Arbor, Michigan 48109,
United States, (2) College of Literature,
Science, and the Arts, University
of Michigan, Ann Arbor, Michigan
48109, United States, (3) Office of
Clinical Pharmacology, Food and Drug
Administration, Silver Spring, Maryland
20993, United States

Adverse drug reactions (ADRs) are
undesirable effects associated with a
clinical use of drugs originating from
their pharmacological on- and off-target
actions. ADRs often prevent patients from
receiving life-saving therapies and can
severely impair a patient’s quality of life.
Neuropathy occurs to approximately 15%
of all FDA-approved drugs in variable
degrees. Despite this high prevalence,
there are no published studies identifying
the common biological signatures of these
neuropathy-causing drugs. The purpose of
this study was to develop a comprehensive
pharmacological network of drugs, known
drug-targets, metabolizing enzymes or
transporters, and potential mediators, all
of which are associated with neuropathy.
Multiple existing and novel bio- and
chem-informatics tools were employed to
efficiently collect relevant chemical and
pharmacological information from multiple
knowledge-bases. A text-mining approach
was used to identify neuropathy-causing
drugs from FDA drug labels (Drugs@FDA
and DailyMed) and ADR reports
(SIDER). SciMiner, a literature-mining tool,
was used to survey all relevant PubMed
abstracts and available medical and
pharmacology reviews in Drugs@FDA.
This approach allowed for identification of
clinical information such as severity and
frequency of neuropathy for each drug.
PubChemSR was used to search and
retrieve relevant chemical information
from PubChem and ChemSpider. Drug-
targets and metabolizing enzymes and
transporters of each neuropathy-causing
drug, constituting the backbone of the
pharmacological network, were collected
from multiple databases, including
DrugBank, Therapeutic Target Database
(TTD), and PharmGKB. Additionally,
protein and genomic interaction data
collected from BioGRID and neuropathy-related genes from OMIM were incorporated to extend the network. A preliminary analysis of this comprehensive network suggests that neuron projection-associated genes are significantly affected by these drugs, although they are not direct targets. In conclusion, integration of multiple bio- and chem-informatics approaches enabled construction of a comprehensive pharmacological network of drugs commonly causing neuropathy and identified novel pathways that may underlie the pathogenesis of the drug-induced neuropathy.

3:35 – Intermission

3:50 – 26 – Molecular scaffolds are special and useful guides to discovery. Jeremy J Yang (1,2), jjyang@salud.unm.edu, Cristian G Bologa (1), David J Wild (2), Tudor I Oprea (1).

(1) Department of Internal Medicine, University of New Mexico, Albuquerque, NM 87131, United States, (2) School of Informatics and Computing, Indiana University, Bloomington, IN 47405, United States

Compounds can be associated by their shared molecular scaffolds, where a scaffold consists of one or more ring-systems joined by linkers. These scaffold associations are meaningful and useful in organic chemistry, medicinal chemistry and chemical biology. The importance of scaffolds is understood to derive both from (1) physical associations with compound properties including bioactivity, and (2) artifactual and other human-based associations such as synthesis/optimization strategies (analog series). Cheminformatics methods and tools have existed to analyze and process scaffolds, and there is extensive literature relating to scaffold based methodology, but also problems, including lack of a rigorous definition of scaffold. Given the strong consensus that scaffolds are important, and the prevalence of scaffold based approaches, the lack of standards for scaffold analysis in cheminformatics is notable. We describe a set of scaffold analysis tools and methods developed by our group, available via open-source project UNM-biocomp-Hscaf, and examples of their use applied to the (1) CARLSBAD database, and (2) the NIH Molecular Libraries Program system BARD (BioAssay Research Database). The scaffold analysis algorithm by Wilkins et al. was implemented and extended. We show that scaffold associations can reveal patterns of bioactivity and promiscuity. Also importantly, scaffold based patterns are inherently comprehensible to chemists, thereby facilitating hypothesis generation. Thus, scaffold analysis is a powerful cheminformatics approach, a kind of chemical indexing, which can enable scientists to navigate biological space, and facilitate knowledge discovery in realms such as chemical biology and drug discovery.


School of Informatics and Computing, Indiana University, Bloomington, IN 47405, United States

Being one of common posttranslational modifications (PTMs) of proteins, glycosylation is involved in important biological processes and associated with diseases. Glycan profiling aims at identifying glycan compositions and determining their abundances within
a complex sample by using mass spectrometry. Traditionally, identification of glycans is achieved using MALDI-MS. Here, we present several bioinformatics approaches to assisting automatic glycan profiling from LC-MS experiments by inducing multiple adducts into the sample. As Liquid Chromatography (LC) provides a highly reproducible separation of glycans, we devise algorithmic techniques to utilize the elution order of glycans to eliminate false glycan annotation. We will also present a glycan sequencing algorithm that incorporates glycan fragment ions to further improve glycan identification. Glycans released from both standard and pooled serum glycoproteins are reduced and permethylated prior to the analysis on a LTQ Orbitrap Velos hybrid FT-MS. A pre-defined library of glycan compositions and adducts masses are used to annotate putative glycan ions observed in the experiments. We used the abduct ion profile of each glycan, consisting of ions formed by the glycan with different adducts, to enhance the confidence of our annotation. We also implemented a longest common subsequence algorithm to match the observed elution order of annotated glycans with their expected order to further eliminate false glycan annotations. We combine these approaches with a glycan sequencing algorithm (when MS/MS data is available) to achieve a high confident glycan annotation in LC-MS data.

4:50 – 28 – Probabilistic inference of reactions and ligands in human gut microbial communities from metagenomic sequences. Dazhi Jiao, djiao@indiana.edu, David J Wild.

Indiana University, Bloomington, In 47405, United States

Shotgun metagenomics has been applied to the studies of the functionality of various microbial communities. As a critical analysis step in these studies, biological pathways are reconstructed based on the genes predicted from metagenomic shotgun sequences. Pathway reconstruction provides insights into the functionality of a microbial community and can be used for comparing multiple microbial communities. Using a probabilistic approach by sampling functions of catalytically promiscuous enzymes within the context of the entire metabolic networks reconstructed from the annotated metagenomes, we analyze metabolic reactions and ligands in human gut microbial communities that were part of the the Human Microbiome Project (HMP) and Metagenomics of the Human Intestinal Tract (MetaHIT) project. Our results show a diversified spectrum of different metabolic reactions and ligands the human gut microbial communities.

5:20 – Concluding Remarks

Sunday, September 8, 2013

CINF Scholarship for Scientific Excellence – EVE Session

Indiana Convention Center
Room: Wabash Birm 1/2

Guenter Grethe, Organizers

6:30 pm – 8:30 pm

29 – Enhanced ranking of PknB inhibitors using data fusion methods. Abhik Seal, abseal@indiana.edu, David John Wild.

School of Informatics and Computing, Indiana University, bloomington, INDIANA 47408, United States

Mycobacterium tuberculosis encodes 11
putative serine-threonine proteins Kinases (STPK) which regulates transcription, cell development and interaction with the host cells. From the 11 STPKs three kinases namely PknA, PknB and PknG have been related to the mycobacterial growth. From previous studies it has been observed that PknB is essential for mycobacterial growth and expressed during log phase of the growth and phosphorylates substrates involved in peptidoglycan biosynthesis. In recent years many high affinity inhibitors are reported for PknB. This paper describes about how data fusion algorithms can identify top PknB inhibitors with high affinity. It has been observed that previous implementation of data fusion has shown effective enrichment of active compounds in both structure and ligand based approaches.

In this study we have used three types of data fusion ranking algorithms on the PknB dataset namely, sum rank, sum score and reciprocal rank. We have identified reciprocal rank algorithm is capable enough to select high affinity compounds earlier in a virtual screening process. Specifically, we found that the ranking of Pharmacophore search, ROCS and Glide XP fused with a reciprocal ranking algorithm not only outperforms structure and ligand based approaches but also capable of ranking actives better than the other two data fusion methods using the BEDROC, robust initial enhancement (RIE) and AUC metrics. We have also screened the Asinex database with best performing reciprocal rank algorithm to identify possible inhibitors for PknB. Using PCA we have shown that the predicted 45 compounds maps well to the PknB Inhibitor chemical space and can be further taken for experimental validation.

30 – 3D-QSAR using quantum-mechanics-based molecular interaction fields.

Ahmed El Kerdawy (1), ahmed.elkerdawy@chemie.uni-erlangen.de, Stefan Güssregen (2), Hans Matter (2), Matthias Hennemann (1,3), Timothy Clark (1,3,4).

(1) Computer-Chemistry-Center, Friedrich-Alexander University Erlangen-Nurnberg, Erlangen, Bavaria 91052, Germany, (2) R&D, LGCR, Structure, Design and Informatics, Sanofi-Aventis Deutschland GmbH, Frankfurt am Main, Germany, (3) Interdisciplinary Center for Molecular Materials, Friedrich-Alexander University Erlangen-Nurnberg, Erlangen, Bavaria 91052, Germany, (4) Centre for Molecular Design, University of Portsmouth, Portsmouth, United Kingdom

The natural evolution of the different computer-aided drug design (CADD) methods involves a shift toward using quantum-mechanics (QM)-based approaches. This shift is not only the result of the ever growing computational power but also due to the need for more accurate and more informative approaches to describe molecular properties and binding characteristics than the currently available ones. QM approaches do not suffer from the limitations inherent to the ball-and-spring description and the fixed atom-centered charge approximation in the classical force fields mostly used by CADD methods. In this project we introduce a protocol for shifting 3D-QSAR, one of the most widely used ligand-based drug design approaches, through using QM based molecular interaction fields (MIFs) which are the electron density ($\rho$), hydrogen bond donor field (HDF), hydrogen bond acceptor field (HAF) and molecular lipophilicity potential (MLP) to overcome the limitations of the current force-field-based MIFs. The average performance of the QM-MIFs (QMFA) models for nine data sets was found to be...
better than that of the conventional force-field-based MIFs models. In the individual data sets, the QMFA models always perform better than, or as well as, the conventional approaches. It is particularly encouraging that the relative performance of the QMFA models improves in the external validation.

31 – Harvard Clean Energy Project: From big data and cheminformatics to the rational design of molecular OPV materials. Johannes Hachmann, jh@chemistry.harvard.edu, Roberto Olivares-Amaya, Alan Aspuru-Guzik.

Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, United States

The Harvard Clean Energy Project (http://cleanenergy.harvard.edu) is a computational effort for the discovery and development of new materials for plastic solar cells in the spirit of the Materials Genome Initiative. A virtual high-throughput infrastructure allows us to characterize millions of material candidates using first principles quantum chemistry. So far, we have studied an unprecedented 2.3 million molecular motifs in 150 million DFT calculations, accumulating 400TB of data. This high-level work is complemented with techniques adopted from drug discovery, cheminformatics, pattern recognition, and machine learning. The challenge of any big data project is to give meaning to such a volume of information. This poster will in detail discuss our data access, mining, and analysis approaches. Our vast collection of computational and experimental results is compiled in a reference database (CEPDB), which is designed as a public information hub for the organic electronics community (similar to the PDB in biology). It allows our collaborators to readily identify candidates with specific property combinations. This on-demand access to structures with any desired set of parameters makes it an ideal tool to find tailored candidates for the different requirements of organic electronic applications. The calibrated in silico results can be used as input for physical derived or empirically generated performance models. CEPDB also allows us to study global trends using the results in their entirety. Our extensive data collection provides a unique foundation for the study of the underlying structure-property relations and OPV design rules. The gained insights can open the door to a rational, systematic, and accelerated development of future high-performance materials. A further task of CEPDB is to supply benchmarks for the performance of the theoretical methods employed in the field (comparable to the more general but much smaller NIST CCCBDB), a testbed for algorithms, as well as a parameter repository for semiempirical or model hamiltonians.

Monday, September 9, 2013

Integrative Chemogenomics Knowledge Mining Using NIH Open Access Resources – AM Session

Indiana Convention Center
Room: 140

Rajarshi Guha, Tudor Oprea, Paul Clemons, Organizers
Rajarshi Guha, Presiding
8:50 am – 12:00 pm

8:50 – Introductory Remarks

8:55 – 32 – Pushing chemical biology data through the pipes: Architecting and extending the BARD API. **Rajarshi Guha**, rajarshi.guha@gmail.com, John Braisted, Ajit Jadhav, Dac–Trung Nyguen, Tyler Peryea, Noel Southall.

*Informatics, National Center for Advancing Translational Sciences, Rockville, MD 20850, United States*

The BioAssay Research Database (BARD) was conceived as a next generation chemical biology resource that would be the go-to public small molecule science resource for translational research, capable of informing the entire small molecule discovery and development process. The BARD system consists of multiple connected components and in this talk we will focus on the BARD API, a RESTful interface that supports programmatic access to the rich datasets contained within the BARD data warehouse, and the basis for multiple BARD client interfaces. We will first give a broad overview of the backend database system and support infrastructure, followed by a discussion of the API design and its capabilities. A key feature of the BARD API is its extensibility via plugins. A BARD plugin appears to users as a native part of the API and can provide novel features ranging from external database access to predictive models to API users. The flexibility of the plugin systems allows plugin authors to provide multiple, arbitrary complex interfaces in front of their plugin - ranging from a JSON response to a HTML5 based rich interface. We will discuss the requirements for a BARD plugin and briefly describe the structure of some exemplar plugins.

9:20 – 33 – QSAR modeling on the web. **Diane Pozefsky** (2), pozefsky@cs.unc.edu, Diptorup Deb (2), Chi Xie (2), Alexander Sedykh (1), Alex Tropsha (1).

(1) School of Pharmacy, University of North Carolina, Chapel Hill, NC 27514, United States, (2) Department of Computer Science, University of North Carolina, Chapel Hill, NC 27955, United States

The use of QSAR technologies is pervasive in most cheminformatics areas. Unfortunately, published QSAR models are rarely provided in an easy to reuse format. Guided by our experience with the ChemBench platform (chembench.mml.unc.edu) and utilizing pluginAPI capabilities offered by the NIH Bioassay Research Database (BARD) framework, we have deployed QSAR model development and distribution option in an open-science, open-access format. We will describe the details of deploying the QSAR modeling capability (using freely available CDK descriptors and publicly available machine learning approaches) as a BARD plugin and highlight the challenges in building a predictive model within BARD. We then describe a few simple usages of QSAR module as a BARD plugin. The use of BARD for model deployment in close proximity to the chemogenomic data stored in BARD enables the community to gain access to novel predictive methods in a uniform fashion, leading to enhanced reuse and improved reproducibility.

9:45 – 34 – Data curation and formal BioAssay Ontology (BAO)-based annotations of the DrugMatrix enable bioactivity-based target-relationship analysis and demonstrate incorporation of external datasets into the BioAssay...
Research Database (BARD). Tudor Oprea (1), toprea@salud.unm.edu, Stephan Schurer (2,3), Ahsan Mir (3), Uma Vempati (3), Jeremy Yang (1), Oleg Ursu (1), Christian Bologa (1).

(1) Translational Informatics Division, Department of Internal Medicine, University of New Mexico, Alberqueque, NM MSC09 5025, United States, (2) Department of Pharmacology, University of Miami, Miami, FL, United States, (3) Center for Computational Science, University of Miami, Miami, FL, United States

The development of a translational informatics knowledgebase (TIK) that is assay-centric and captures relevant information from chemogenomics, proteomics and phenomics requires appropriate annotations for each of the classes of objects within the TIK system. Careful curation is required at all levels, including chemical, biological assay with target, phenotype, pathway, detection, etc. - aspects that are both enabled and required within the BARD system. Here we describe the process of annotating and evaluating an external screening set, the bioactivity component of DrugMatrix (baDM). The complete DrugMatrix dataset is freely available at https://ntp.niehs.nih.gov/drugmatrix/index.html, offered by the National Toxicology Program (NTP). baDM assays were extracted from ChEMBL (https://www.ebi.ac.uk/chemblpdb/), with additional re-evaluation against the complete NTP set. Assay information was extracted from the original assay provider, Eurofins Panlabs (https://www.eurofinspanlabs.com/Catalog/AssayCatalog/AssayCatalog.aspx?search). The baDM set from ChEMBL contains 871 chemicals, measured in 131 assays, with 7055 exact Ki or IC50 values. BARD leverages BioAssay Ontology (BAO) and BAO annotations including biochemical information (e.g., reference compound structures and bioactivities) were added for all baDM assays. Target curation was required for 37 proteins, with complete target reassignment in some cases (e.g., “central imidazoline I2 receptors” are in fact an allosteric site of monoamine oxidases A or B). A regression analysis of the complete baDM matrix shows target relationships at the bioactivity level: Highly correlated sets (R^2 >= 0.5) were observed within the adrenergic, dopaminergic, muscarinic, opioid and serotonergic receptor family subtypes, as well as within the calcium channel L-type sites, with up to 87 chemicals.

10:10 – Intermission

10:20 – 35 – Development of the BioAssay Research Database (BARD): A user-friendly perspective based on active participation from biologists and chemists. Eric S Dawson (1,4,6), eric.dawson@vanderbilt.edu, Shaun R Stauffer (2,3,5,6), Craig W Lindsley (2,3,5,6).

(1) Department of Biochemistry, Vanderbilt University, Nashville, Tennessee 37232, United States, (2) Department of Pharmacology, Vanderbilt University, Nashville, Tennessee 37232, United States, (3) Department of Chemistry, Vanderbilt University, Nashville, Tennessee 37232, United States, (4) Center for Structural Biology, Vanderbilt University, Nashville, Tennessee 37232, United States, (5) Center for Neuroscience Drug Discovery, Vanderbilt University, Nashville, Tennessee 37232, United States, (6) Specialized Chemistry Center for Probe Development (MLPCN), Vanderbilt University, Nashville, Tennessee 37232, United States
The Molecular Libraries Program (MLP), an NIH Common Fund Initiative, has recently fostered a unique collaboration among prominent academic centers[1] to develop next-generation, user-friendly tools to provide enhanced access to MLP bioassay data. The MLP network has generated a treasure trove of valuable data that now requires a state-of-the-art bioassay database to enable a broad base of scientists in the larger research community to query, mine, and analyze these data to generate novel scientific hypotheses. Our multidisciplinary collaborative approach allows scientists to annotate all MLP data using a shared language to provide facile access to data while integrating existing chemical biology resources. Early software development contributions to user-interface and workflow design by expert medicinal chemists and biologists with industrial drug discovery experience are enabling meaningful analyses and interpretation of probe-development data without a requirement for users to have expertise in informatics. Intuitive queries that achieve robust connections from broad analysis of experimentally measured biological activity to medicinal chemistry structure activity relationships support hypothesis generation to streamline probe and drug discovery projects. The BioAssay Research Database (BARD) is scheduled for public beta release in June 2013 to meet the needs of chemical biology researchers in a user-driven, flexible, and innovative fashion. [1] Broad Institute, University of Miami, National Center for Advancing Translational Sciences (NCATS), Sanford-Burnham Medical Research Institute, the Scripps Research Institute, University of New Mexico and Vanderbilt University

10:45 – 36 – BADAPPLE promiscuity plugin for BARD: Evidence-based promiscuity scores. Jeremy J Yang (1), jeremyjyang@gmail.com, Oleg Ursu (1), Cristian G Bologa (1), Anna Waller (2), Larry A Sklar (2), Tudor I Oprea (1).

(1) Department of Internal Medicine, University of New Mexico, Albuquerque, NM 87131, United States, (2) Department of Pathology, University of New Mexico, Albuquerque, NM 87131, United States

In chemical biology and drug discovery, promiscuity concepts have grown in importance, along with new conceptual frameworks including systems biology and systems pharmacology. Accordingly, one-dimensional notions of promiscuity have evolved toward more complex approaches, for example conditioned on target classes. Artifactual issues accompanying real-world bioassay data have also been an important consideration, that is, false positives associated with assay methodologies. Thus making effective use of bioassay data requires understanding promiscuity both as a biological phenomena and experimental source of error. Mindful of these new challenges, we have developed a component for BARD (BioAssay Research Database) called the “Badapple Promiscuity Plugin” (BioActivity Data Associative Promiscuity Pattern Learning Engine). The Badapple algorithm generates a score, based on scaffold-family membership, and derived solely from empirical activity data in BARD. The score reflects both a pan-assay “batting average” and the weight of evidence, thus high scores indicate confidence. “Evidence-based” or “data-driven” further implies that the algorithm evaluates data “as is”, thus scores may...
change as new evidence becomes available. Fully integrated with BARD via the flexible IPlugin specification, BARD’s semantic advances provide a valuable and unique synergy. The new annotations and bioassay ontology, based on BAO and adapted and implemented for BARD, enable improvements, extensions, and customizations for Badapple. In this presentation we will describe these tools and some examples from molecular discovery scenarios.

11:30 – Panel Discussion

Monday, September 9, 2013

Role and Value of Social Networking in Advancing the Chemical Sciences – AM Session Social Media for the Individual Scientist and to Support Education

Indiana Convention Center
Room: 141

Cosponsored by CHED, SCHB, YCC

Antony Williams, Jennifer Maclachlan, Organizers
Jennifer Maclachlan, Presiding

8:15 am – 12:00 pm

8:15 – Introductory Remarks

8:20 – @ChemConnector and my personal experiences in participating in the expanding social networks for science. Antony J. Williams, williamsa@rsc.org.

eScience and Cheminformatics, Royal Society of Chemistry, Wake Forest, NC 27587, United States

The number of social networking sites available to scientists continues to grow. We are being indexed and exposed on the internet via our publications, presentations and data. We have many ways to contribute, annotate and curate, many of them as part of a growing crowdsourcing network. As one of the founders of the online ChemSpider database I was drawn into the world of social networking to participate in the discussions that were underway regarding our developing resource. As a result of my experiences in blogging, and as a result of developing collaborations and engagement with a large community of scientists, I have become very immersed in the expanding social networks for science. This presentation will provide an overview of the various types of networking and collaborative sites available to scientists and ways that I expose my scientific activities online. Many of these activities will ultimately contribute to the developing measures of me as a scientist as identified in the new world of alternative metrics.

8:45 – 38 – Get to the point! Techniques for organizing your online social networking skills. Tom Ruginis, happilabs@gmail.com.

HappiLabs, LLC, Chicago, IL 60657, United States

It’s easy to assume that social media can be the catalyst for getting noticed, growing your business, or finding happiness. Assume nothing. If used incorrectly, social media can harm your image and waste your time &money. However, if used correctly, you’ll create interactions that expose you to a new world of opportunities—new colleagues, new customers, and new friends. Learn how to use social media effectively and efficiently by getting to the point. This talk will provide techniques to help you navigate common obstacles that people and businesses
are exposed to, such as: What do I say and how do I say it? How do I efficiently manage my time online? At the end of the day, am I getting through to people or am I annoying them? We’ll discuss all of these questions and provide you with solutions based on published reports and Tom’s experience as a social media specialist. About the speaker: Tom Ruginis was formerly a Molecular Biology PhD student turned Marketing Specialist for a lab supply distributor. He now helps scientists communicate with two companies he has founded: RuugyMedia, a social media consulting company for science and environmental organizations, and HappiLabs.org, a research organization that studies scientists.


Plum Analytics, Philadelphia, PA 19025, United States

For decades the measurement of the influence of experts and scholars has been based on an author publishing papers in prestigious (print) journals, and having subsequent experts formally cite that researcher’s work in the same or other prestigious (print) journals. Today, however, scholarly communication is conducted online and in multiple digital formats that leave a trail of data exhaust. Plum Analytics gathers this scholarly exhaust, quantifies it, and reports on the people, labs, corporations, and research institutes who have the most impact. This gives researchers, and those supporting and funding them, an edge in an increasingly competitive world. Metrics are available from a wide variety of sources and can be aggregated to give alternative ways of measuring the impact of research. Dubbed “altmetrics,” this growing field is providing new ways of looking at traditional measurements of engagement and interaction in scholarly communication. Andrea Michalek, co-founder of Plum Analytics, will discuss the new metrics that are increasingly useful in evaluating the impact of current research in the chemical sciences, including the types of metrics that are available, how they are used in determining impact, and developments at Plum Analytics.

9:35 – Intermission

9:45 – 40 – Using social tools collaboratively to communicate and advance science. Christopher McCarthy, c_mccarthy@acs.org, Christine Brennan-Schmidt, c_schmidt@acs.org.

American Chemical Society, Washington, District of Columbia 20036, United States

In the last decade, the number of social tools available to communicate and engage communities has increased dramatically. These tools have important applications for advancing chemistry, especially to promote science literacy and when partnered with in-person community outreach activities. This presentation will explore social tools, including Facebook, Twitter, and the ACS Network, and, through the use of case studies, offer suggestion for using these tools together to communicate with a variety of audiences.

10:10 – 41 – Social media in cheminformatics education. David J Wild (1), djwild@indiana.edu, Robert Belford (2).
At Indiana University School of Informatics and Computing, we have over the last decade developed a graduate teaching curriculum in cheminformatics, with a strong emphasis on enabling remote participation by students using distance learning and social media technologies. We have created a collection of free wiki-based resources, and are expanding this with videos, links to external resources, and a low cost introductory cheminformatics eBook (see http://icep.wikispaces.com). In the last year, we began a joint NSF-funded project with the University of Arkansas at Little Rock, to create a Cheminformatics OLCC, a hybrid online and local learning environment to enable chemistry undergraduates in distributed locations to learn the basics of cheminformatics and their application in chemistry. This project is based heavily on social media, including forums and tagging, to maximize student and facilitator involvement and exchange. This talk will review the technologies and resources that have been, and are being, developed and utilized, and the opportunities they afford, as well as the challenges, will be discussed.

10:35 – 42 – ChemConf to ConfChem: Twenty years and counting. Robert E. Belford (1), rebelford@ualr.edu, Harry E. Pence (2).

This year we are celebrating 20 years of the online ConfChem conference. The first ConfChem (ChemConf) predated the World Wide Web, with papers distributed as ASC II text files over gopher or ftp servers, and discussed over the ConfChem listserv. The second ConfChem used web 1.0 technologies (complete with instructions on the use of Netscape), while current ConfChems utilized web 2.0 technologies. In the world of the internet there are few online entities that have survived two decades of the changes induced by evolving technologies. Yet throughout this time the core ConfChem model of internet mediated social interaction has essentially remained a constant. What can we learn from two decades of ConfChem? With an eye on the future, we will look at both the evolution of ConfChem, and the constants that have enabled it to survive the ephemeral landscape of the internet.

11:00 – Intermission

11:10 – 43 – Small chemical businesses and the importance of being social. Jennifer L Maclachlan, pidgirl@gmail.com.

PID Analyzers, LLC, Sandwich, Massachusetts, MA 02563, United States

As a small chemical business owner I utilize the following social media platforms for my business: Facebook, Twitter, Hootsuite, LinkedIn, Blogger, Tumblr, Google+, Constant Contact and Pinterest. I will discuss the integral role that each of these social media vehicles plays in the sales, marketing, and branding of my analytical instrumentation. Additionally I will speak to how I’ve been able to maintain and grow existing business relationships through consistent social media communications as well as continuously building a following. Metrics
for tracking leads to social campaigns will be addressed. Lastly I will discuss the importance of finding your social media voice and getting out there and being social.

11:35 – 44 – XCITR (Explore Chemical Information Teaching Resources): A community resource for chemistry instructors, chemistry librarians, and chemical information providers. 
Grace Baysinger (1), graceb@stanford.edu, Andrea Twiss-Brooks (2), Guenter Grethe (3), Gregor Fels (4).

(1) Stanford University, United States, (2) University of Chicago, United States, (3) Unaffiliated, United States, (4) University of Paderborn, Germany

XCITR (http://www.xcitr.org) is an international collaborative effort to provide a hub for sharing chemical information teaching resources more effectively and efficiently. While the community and the collection are still growing, XCITR already contains a vibrant and diverse mix of people and resources. Instead of “reinventing the wheel,” resources can be reused, adapted for local use, or serve as inspiration for developing new teaching resources. This presentation will summarize the current state and future directions for XCITR.

Monday, September 9, 2013

Science-Based Policy Development in the Environment, Food, Health, and Transport Sectors – PM Session

Indiana Convention Center
Room: 140

Cosponsored by AGFD, ANYL, ENVR, MEDI

William Town, Organizers
William Town, Presiding

1:15 pm – 5:05 pm

1:15 – Introductory Remarks

1:20 – 45 – Adaptive management tools for engineered nanomaterials in municipal wastewater effluents. Thomas A. Duster, tduster@nd.edu.

Department of Civil and Environmental Engineering and Earth Science, University of Notre Dame, Notre Dame, IN 46556, United States

The ubiquity of engineered nanomaterials in consumer products results in their delivery to municipal wastewater treatment systems, where they may be subsequently discharged to the environment. At sufficient concentrations, many common nanomaterials, including titanium dioxide nanoparticles and carbon nanotubes, are toxic or disruptive to aquatic organisms, but significant challenges exist regarding the application of contemporary environmental policies to mitigate these potential impacts. For example, the traditional standards-to-permits approach of the Clean Water Act (CWA), which applies to most wastewater treatment plant effluents in the United States, typically involves the development of contaminant-specific water quality criteria. However, existing research regarding the detection, fate, and toxicology of nanomaterials is still in its infancy and rapidly changing, thereby limiting the ability of policymakers to justify and establish static effluent discharge standards for these emerging contaminants. Hence, I describe herein an adaptive nanomaterial management approach that strives to bridge the gap.
between significant scientific uncertainties and an ostensive need for some type of policy structure. At the core of this adaptive management procedure is a robust mechanism for information and data organization, which is programed to alert policymakers of convergence in the literature between: (a) observed and/or anticipated concentrations of target nanomaterials in wastewater effluents; (b) demonstrated impacts of these concentrations on aquatic organisms or ecological function; and (c) our technological capacity to reliably detect these target nanomaterial concentrations. The confluence of these factors is expected to be a significant trigger in evaluating the need for specific management actions and/or expansion of policies related to the release of engineered nanomaterials to environmental systems. Finally, I describe how specific elements of this approach may be applied to policy challenges for other emerging contaminants.

2:20 – 47 – Identification of pathways of toxicity to predict human effects. Helena T Hogberg, hhogberg@jhsp.edu, Thomas Hartung.

Department of EHS, The Johns Hopkins University, Bloomberg School of Public Health, Center for Alternative to Animal Testing (CAAT), Baltimore, Maryland 21205, United States

The National Research Council report from 2007 “Toxicity Testing in the 21st Century: A vision and a strategy” has created an atmosphere of departure in the US. It suggests moving away from traditional (animal) testing to modern technologies based on pathways of toxicity. These pathways of toxicity could be modeled in relatively simple cell tests. The NIH is funding, by a transformative research grant, The Human Toxome project led by CAAT. The project involves US EPA ToxCast, Hamner Institute, Agilent and members of the Tox-21c panel. The goal is to develop a public database of pathways, the Human Toxome, to enable scientific collaboration and exchange. An area of toxicity where Tox-21c could have significant impact is developmental
neurotoxicity (DNT). Current animal tests for DNT have several limitations: high costs ($1.4 million per substance) and time consumption. In addition, there are scientific concerns regarding the relevance of these studies for human health effects. Consequently, only few substances have been identified as developmental neurotoxicants. This is a concern as evidence shows that exposures to environmental chemicals contribute to the increasing incidence of neurodevelopmental disorders in children. Moving towards a mechanistic science can help us identify the perturbed pathways that likely lead to these adverse effects. DNTox-21c is a CAAT project funded by FDA that aims to identify pathways of developmental neurotoxicity using a metabolomics approach. Beside the technical development of new approaches, a case is made that we need both conceptual steering and an objective assessment of current practices by evidence-based toxicology. It is suggested to apply an approach modeled on Evidence-based Medicine (EBM), which over the last two decades has demonstrated that rigorous systematic reviews of current practices of studies provide powerful tools to provide health care professionals and patients with the current best scientific evidence for diagnostic and treatment options.


Policy divergence in the absence of science: the case of e-cigarettes
Over the past five years electronic cigarettes (e-cigarettes) have emerged as a new consumer product that is being used by an increasing number of smokers who are seeking less risky alternatives to conventional cigarettes. E-cigarettes tend to be designed to look and feel similar to conventional cigarettes, but they do not contain tobacco. They are battery-powered devices that produce an aerosol usually containing nicotine. Currently, there is significant inconsistency
in the way that e-cigarettes are being regulated around the world: e-cigarettes are banned in some countries or are being regulated either as medicinal, tobacco or general consumer products in others. There is also a diversity of views regarding the potential role that e-cigarettes could play in helping to reduce the public health impacts of tobacco use. In fact, the science to support this emerging category of products is only being developed now, and there are many gaps. E-cigarettes therefore represent a timely case study on what can happen as regards policy development for regulation of a new product category in the absence of a solid scientific foundation. Some views will also be presented on how the development of such a scientific foundation might be accelerated to, in turn, help inform development of an appropriate regulatory framework for e-cigarettes.

4:05 – 50 – Role of regulatory science in reducing the public health impact of tobacco use. Christopher J Proctor, christopher_proctor@bat.com, Chuan Liu.

Group Research & Development, British American Tobacco, Southampton, United Kingdom SO15 8TL, United Kingdom

The US FDA, through the 2009 US Family Smoking and Prevention Tobacco Control Act, are introducing a variety of regulations aimed at reducing the public health impact of tobacco use. These include considering the levels of harmful and potentially harmful constituents of tobacco products and regulations governing modified risk tobacco products. FDA have set out a series of research questions that they believe are research needed to underpin their regulatory proposals and have initiated a large research funding programme, in association with NIH. Other scientific advisory groups, including the World Health Organisation’s Scientific Advisory Committee on Tobacco Product Regulation have also listed research needed to assist the development of science-based public policy on tobacco. This presentation will summarise the research questions being framed by regulators as related to product regulation and provide some views on how the development of regulatory science in tobacco might be accelerated.

4:35 – 51 – Systematic and structural risk analysis approaches for establishing maximum levels of essential nutrients and other bioactive substances in fortified foods and food supplements. David P Richardson, info@dprnutrition.com.

School of Chemistry, Food and Pharmacy, University of Reading, Reading, United Kingdom

Nutritional risk analysis addresses the essential nutrients and other substances with nutritional and physiological effects and the risk to health from their inadequate and/or excessive intake. The paper reviews the principles of risk management in order to underpin regulatory developments around the world to establish maximum amounts of vitamins and minerals and other substances in fortified foods and food supplements. The proposed science-based risk management models for public health decision-making take into account international risk assessments and (1) the tolerable upper intake levels (ULs) for vitamins and minerals, (2) the highest observed intakes (HOIs) for bioactive substances for which no adverse effects have been identified, and (3) the contributions to total intake from conventional foods, fortified foods and food supplements. The models propose the allocation of nutrient substances into three categories of risk and maximum
levels in order to protect consumers, both adults and children, from excessive intakes.

Monday, September 9, 2013

Role and Value of Social Networking in Advancing the Chemical Sciences – PM Session Social Media to Share Science with the Community and the Sharing of Chemical Information

Indiana Convention Center
Room: 141

Cosponsored by CHED, SCHB, YCC

Antony Williams, Jennifer Maclachlan, Organizers
Antony Williams, Presiding

1:00 pm – 3:45 pm

1:00 – 52 – Exploiting the digital landscape to advance the chemical sciences. Bibiana Campos-Seijo, camposseijob@rsc.org, Will Russell, Vibhuti Patel.

Royal Society of Chemistry, Cambridge, United Kingdom

The new digital landscape has resulted in information overload on the web. Making content available is no longer enough: you have to ensure it is also discoverable. A key aspect of this is social media, which since 2007 has become the established norm and is no longer just a way for teenagers to interact. Digital Natives (Generation Y) intuitively utilise the correct tools for discovering and sharing content online: social media and traditional media are merging.

This presentation will look at the huge take-up of digital technologies and the changes they are bringing to the publishing and discovery models. Chemistry World, the flagship magazine of the Royal Society of Chemistry, provides us with the perfect subject case for a discussion of how a social media strategy can successfully be implemented in a traditional publishing environment. With more than 200,000 followers on Twitter and 50,000 on Facebook, we’ll demonstrate that social media tools are powerful levers to improve discoverability and engagement, promote the dissemination of scientific information, facilitate networking and the establishment of collaborations, and ultimately, to advance the chemical sciences.

1:25 – 53 – Using social media as a way to communicate science to the community. George Ruger, gruger04@yahoo.com.

ACS Mid Hudson, Modena, NY 12548, United States

Social Media is very useful to help communicate the sciences to the scientific community as well as to the general population. Blogs and other sources can help advertise events beforehand and can also help to share the information after the event is over. Examples of successful advertisements and post-event wrap-ups will be given.

1:50 – 54 – Collaborating to convey chemical knowledge through Wikipedia. Martin A Walker, walkerma@potsdam.edu.

Department of Chemistry, State University of New York at Potsdam, Potsdam, New York 13676, United States

Wikipedia provides scientists and the general public with a vast array of chemical information. This presentation will examine the ways this resource was built, and how past and present collaborations continue.
to add value. The role of social networking will be discussed, including the challenges and benefits of crowdsourced information.

2:15 – Intermission

2:25 – 55 – Tools and strategies for sharing chemical information and research on the web. Jean-Claude Bradley (1), bradlejc@drexel.edu, Andrew Lang (2).

(1) Department of Chemistry, Drexel University, Philadelphia, PA 19104, United States, (2) Department of Mathematics, Oral Roberts University, Tulsa, OK 74171, United States

This presentation will outline effective strategies and specific tools for leveraging student work in both classroom and research environments for contributing to the advancement of scientific knowledge in practical and open ways. Examples from organic chemistry, cheminformatics and modeling will be discussed. The importance of parallel communications channels optimizing for human or machine readability will be stressed. It will be demonstrated that openness is necessary for the efficient flow of information within this evolving system. Finally, the evaluation of the impact of such projects and the possibility of its quantification will be explored.


National Center for Biotechnology Information, National Institutes of Health, Bethesda, MD 20894, United States

Social networking and the ability to share information with your colleagues or friends has become commonplace using platforms such as FaceBook, Twitter, and blogs. The adoption of social networking in the chemical information space is growing as a forum to share and discuss scientific thought. This talk will provide an overview and highlight social networking resources and forums provided by PubChem and NCBI.

3:15 – Panel Discussion

3:40 – Concluding Remarks

Monday, September 9, 2013

Joint CINF-CSA Trust Symposium: Semantic Technologies in Translational Medicine and Drug Discovery – PM Session

Indiana Convention Center
Room: 142

David Wild, Jan Kuras, Organizers
David Wild, Jan Kuras, Presiding

1:30 pm – 5:25 pm

1:30 – Introductory Remarks

1:35 – 57 – Building support for the semantic web for chemistry at the Royal Society of Chemistry. Valery Tkachenko (1), tkachenkov@rsc.org, Colin Batchelor (2), Jon Steele (2), Alexey Pshenichnov (1), Antony J. Williams (1).

(1) eScience and Cheminformatics, Royal Society of Chemistry, Wake Forest, NC 27587, United States, (2) eScience and Cheminformatics, Royal Society of Chemistry, Cambridge, United Kingdom

The Royal Society of Chemistry provides a variety of databases and services covering multiple domains of Chemistry. That includes our electronic publishing platform,
ChemSpider and its related databases, the National Chemistry Database and digital access to the RSC archive that spans over 170 years. In order to support the rising tide of semantic web technologies we are now working on exposing our data to conform with the linked data paradigm. This presentation will provide an overview of our work to introduce semantic structure to all RSC electronic resources as well as outlining ways to access this information using standard formats and various APIs.

Evan Bolton, bolton@ncbi.nlm.nih.gov, Gang Fu, Paul Thiessen, Asta Gindulyte.

National Center for Biotechnology Information, National Institutes of Health, Bethesda, MD 20894, United States

PubChem is a community driven chemical biology resource containing information about the biological activities of small molecules. With over 200 contributors, PubChem is a sizeable resource with over 115 million sample descriptions, 46 million unique small molecules, and 200 million biological activity results. Resource description framework (RDF) is a family of World Wide Web Consortium (W3C) specifications used as a general method for concept description. The RDF data model can encode semantic descriptions in so-called triples (subject-predicate-object). This talk will give an overview of the PubChemRDF project scope and some examples of its use.

Gang Fu (1), gang.fu@nih.gov, Colin Batchelor (2), Michel Dumontier (3), Janna Hastings (4,5), Hande Küçük (6), Stephan Schurer (6,7), Uma Vempati (6), Egon Willighagen (8), Evan Bolton (1).

(1) National Center for Biotechnology Information, National Library of Medicine, National Institute of Health, Bethesda, MD 20894, United States, (2) Royal Society of Chemistry, Cambridge, United Kingdom, (3) Department of Biology, Institute of Biochemistry, School of Computer Science, Carleton University, Ottawa, Canada, (4) European Bioinformatics Institute, Cambridge, United Kingdom, (5) University of Geneva, Genève, Switzerland, (6) Center for Computational Science, University of Miami, Miami, FL, United States, (7) Department of Molecular and Cellular Pharmacology, University of Miami, Miami, FL, United States, (8) Department of Bioinformatics – BiGCaT, Maastricht University, Maastricht, The Netherlands

PubChem serves as an open repository for chemical biology: integrating information on chemical structures, their biological activities, and biomedical annotations. Semantic Web technologies and standards can provide a means for the large-scale integration and reasoning of PubChem data, and may help PubChem data to be shared, reused, and analyzed across chemical, biological, and life science domains. A set of formal ontologies enhancing data integration and interoperability are utilized to encapsulate PubChem domain-specific knowledge, including CHEMical INFormation ontology (CHEMINF), Semanticscience Integrated Ontology (SIO), Chemical Entities of Biological Interest (ChEBI), BioAssay Ontology (BAO), and Gene Ontology (GO). While great care is taken to integrate PubChem concepts within appropriate semantic categories, it is not trivial to do so.

This talk will give an overview of how the PubChem data is integrated to various
ontologies and the opportunities they afford.

3:05 – Intermission

3:20 – 60 – Practical semantics in the pharmaceutical industry: The Open PHACTS project. Antony J. Williams, williamsa@rsc.org.

eScience and Cheminformatics, Royal Society of Chemistry, Wake Forest, NC 27587, United States

The information revolution has transformed many business sectors over the last decade and the pharmaceutical industry is no exception. Developments in scientific and information technologies have unleashed an avalanche of content on research scientists who are struggling to access and filter this in an efficient manner. Furthermore, this domain has traditionally suffered from a lack of standards in how entities, processes and experimental results are described, leading to difficulties in determining whether results from two different sources can be reliably compared. The need to transform the way the life-science industry uses information has led to new thinking about how companies should work beyond their firewalls. In this talk we will provide an overview of the traditional approaches major pharmaceutical companies have taken to knowledge management and describe the business reasons why pre-competitive, cross-industry and public-private partnerships have gained much traction in recent years. We will consider the scientific challenges concerning the integration of biomedical knowledge, highlighting the complexities in representing everyday scientific objects in computerised form. This leads us to discuss how the semantic web might lead us to a long-overdue solution. The talk will be illustrated by focusing on the EU-Open PHACTS initiative (openphacts.org), established to provide a unique public-private infrastructure for pharmaceutical discovery. The aims of this work will be described and how technologies such as just-in-time identity resolution, nanopublication and interactive visualisations are helping to build a powerful software platform designed to appeal to directly to scientific users across the public and private sectors.

3:50 – 61 – Enabling the translational medicine and drug discovery information workflow. David Evans (1), david.evans@reedelsevier.ch, Timothy Hoctor (2), Jacqui Mason (2), Pieder Caduff (1).

(1) Reed Elsevier Properties SA, Neuchatel, Switzerland, (2) Elsevier Inc, New York, NY 10010, United States

Information critical for research decisions often is distribute across disparate content repositories from diverse data providers with diverging nomenclature and in inconsistent formats and often buried in text documents. The creation of consistent, normalized information stores (using taxonomies and ontologies) enables scientists to find appropriate relevant, timely information in order to make the best decisions.

We will describe the development of a number of chemical, pharmacological and target taxonomies and ontologies. We will show the real world application of these to: Build data repositories from large scale text analysis machines Link data repositories from diverse including internal, public & 3rd party sources Enable users to create ontology based concept queries across these repositories.

4:20 – 62 – Standardized drug and
pharmacological class network in ontology representation. Qian Zhu, zhu.qian@mayo.edu, Cui Tao, Christopher Chute.

Department of Health Sciences Research, Mayo Clinic, Rochester, MN 55901, United States

Dozens of drug terminologies and resources capture the drug and/or drug class information, ranging from their coverage and adequacy of representation. No transformative ways are available to link them together in a standard way, which hinders data integration and data representation for drug-related clinical and translational studies. In the meantime, ontology design and representation being widely used in the translational research offers more capabilities to discovery novel associations among entities underneath and consistency checkup. In this paper, we introduce a standardized drug and drug class network construction that integrates multiple drug terminological resources, using Anatomical Therapeutic Chemical (ATC) and National Drug File Reference Terminology (NDF-RT) as network backbone, and expanding with RxNorm and Structured Product Label (SPL). Meta-ontology will be established to capture drug, drug class and drug relevant information, in addition to the relationship among these entities. Instance layer will be built on the top of the meta-ontology.

4:50 – 63 – Application of text mining and semantic technology in external intelligence surveillance. Yiqun H Li, lihy@lilly.com.

Eli Lilly and Company, United States

Pharmaceutical companies are facing increasing pressures from all fronts, rising customer expectations, patent expiration cliff, innovation challenges, complex regulatory environments. In an environment of constrained resources and elevated premiums, it is critical for companies to make the right R&D investments to bring innovations to patients quickly. These drive the need for effective knowledge discovery and management of the exponentially growing public information. Recent advancements of semantic technology show the promise of effective managing and connecting data points in an unprecedented manner. This talk will present an application of text mining and semantics technology in pharmaceutical external intelligence. We will explore the roles of agile text mining, ontology management, semantic data integration in accelerating strategic decision making and R&D operational effectiveness.

5:20 – Concluding Remarks

Monday, September 9, 2013

Sci–Mix – EVE Session

Indiana Convention Center Room: Halls F&G

Jeremy Garritano, Organizers

8:00 pm – 10:00 pm

9 – ChemDraw, iPads, and collaboration tools in the classroom: Results of a joint PerkinElmer and McGraw Hill pilot at the organic chemistry undergraduate level Robin Y Smith (1), robin.smith@perkinelmer.com, Hans C Keil (1), hans.keil@perkinelmer.com, Tamara Hodge (2).

(1) PerkinElmer, Waltham, Massachusetts 02451, United States, (2) McGraw Hill Education, Dubuque, Iowa 52001, United States

In partnership with McGraw Hill Education,
PerkinElmer will conduct a pilot across several undergraduate organic chemistry classes testing the use of iPads, ChemDraw, Chem3D and a new cloud-based collaboration service. The pilot will test the effectiveness of tablet-based learning at the undergraduate level. Working closely with professors and students, PerkinElmer will adapt ChemDraw and other current software for the learning platforms and techniques of tomorrow. McGraw Hill, as a leading company in the education software industry, will work closely with participating professors to analyze the success of the pilot (versus previous classes) and make recommendations on future steps. The pilot will take place over the 2013 summer session with results being compiled in August and presented publically for the first time at the 246th ACS National Meeting.

12 – Enabling chemistry on-the-go with modern solutions. Tamsin E Mansley (1), tamsin.mansley@dotmatics.com, Graeme E Dennis (1), Shikha O’Brien (2).

(1) Dotmatics, Inc., Woburn, MA 01801, United States, (2) Dotmatics, Inc., San Diego, CA 92121, United States

The paradigm in chemical education and research is continuously evolving as technology becomes more pervasive. Even today scientists and students are dependent on a multitude of tools to capture and share data: paper notebooks, calculators, text books, desktop personal computers, etc. Are we playing catch-up to social media and technology when it comes to science? Today laptops, tablets, mobile devices and social media are commonplace and there is a need to provide technology to educators, users and early adopters through these media. This presentation will focus on our ability to support distance learning and on-the-go chemistry in research organizations through web-based and mobile applications, including free tools such as Dotmatics’ Elemental chemical sketching app for iPad, iPhone and Android devices.

15 – Improving access to data: A distributed approach. Graeme E Dennis (1), graeme.dennis@dotmatics.com, Tamsin E Mansley (1), Shikha O’Brien (2).

(1) Dotmatics, Inc., Woburn, MA 01801, United States, (2) Dotmatics, Inc., San Diego, CA 92121, United States

Today’s environment requires decentralized organizations to exchange critical information, often in a variety of formats and through multiple communication channels. Addressing issues of security, data loss and poor communication are essential for any project’s success. Among the greatest challenges faced by the scientists today are (a) dealing with the data deluge and (b) accessing and making sense of the data. How is unstructured data to be handled in a way that permits querying, browsing, and analysis while retaining all the meaning of its original presentation? Strategies permitting scientists to have secure, ready access to their own data, irrespective of where it might be located, in a format that is meaningful to them, will be presented.

22 – OPSIN: Taming the jungle of IUPAC chemical nomenclature. Daniel M Lowe, daniel@nextmovesoftware.com, Peter Murray-Rust, Robert C Glen.

Department of Chemistry, University of Cambridge, Cambridge, Cambridgeshire CB2 1EW, United Kingdom

OPSIN (Open Parser for Systematic IUPAC Nomenclature) is an open source
freely available program for converting chemical names, especially those that are systematic in nature, to chemical structures. The software is available as a Java library, command-line interface and as a web service (opsin.ch.cam.ac.uk). OPSIN accepts names that conform to either IUPAC or CAS nomenclature and can convert them to SMILES, InChI and CML (Chemical Markup Language). OPSIN has grown from covering only simple general organic chemical nomenclature to the point of having competent coverage of all areas of organic chemical nomenclature.

One of the most recent additions is comprehensive support for the nomenclature of carbohydrates. This brings support for dialdoses, diketoses, ketoaldoses, alditols, aldonic acids, uronic acids, aldaric acids, glycosides and oligosaccharides, in both the open chain and cyclic forms, named systematically or from trivial sugar stems with support for modification terms such as anhydro or deoxy. OPSIN’s support for specialised and general organic nomenclature will be demonstrated through illustrative examples and accompanying performance metrics. We focus in particular on areas of nomenclature for which support was recently added and those that are complex to implement such as fused ring nomenclature.

26 – Molecular scaffolds are special and useful guides to discovery. Jeremy J Yang (1,2), jjyang@salud.unm.edu, Cristian G Bologa (1), David J Wild (2), Tudor I Oprea (1).

(1) Department of Internal Medicine, University of New Mexico, Albuquerque, NM 87131, United States, (2) School of Informatics and Computing, Indiana University, Bloomington, IN 47405, United States

Compounds can be associated by their shared molecular scaffolds, where a scaffold consists of one or more ring-systems joined by linkers. These scaffold associations are meaningful and useful in organic chemistry, medicinal chemistry and chemical biology. The importance of scaffolds is understood to derive both from (1) physical associations with compound properties including bioactivity, and (2) artifactual and other human-based associations such as synthesis/optimization strategies (analog series). Cheminformatics methods and tools have existed to analyze and process scaffolds, and there is extensive literature relating to scaffold based methodology, but also problems, including lack of a rigorous definition of scaffold. Given the strong consensus that scaffolds are important, and the prevalence of scaffold based approaches, the lack of standards for scaffold analysis in cheminformatics is notable.

We describe a set of scaffold analysis tools and methods developed by our group, available via open-source project UNM-biocomp-Hscaf, and examples of their use applied to the (1) CARLSBAD database, and (2) the NIH Molecular Libraries Program system BARD (BioAssay Research Database). The scaffold analysis algorithm by Wilkins et al. was implemented and extended. We show that scaffold associations can reveal patterns of bioactivity and promiscuity. Also importantly, scaffold based patterns are inherently comprehensible to chemists, thereby facilitating hypothesis generation. Thus, scaffold analysis is a powerful cheminformatics approach, a kind of chemical indexing, which can enable scientists to navigate biological space, and facilitate knowledge discovery in
realms such as chemical biology and drug discovery.

29 – Enhanced ranking of PknB inhibitors using data fusion methods. Abhik Seal, abseal@indiana.edu, David John Wild.

School of Informatics and Computing, Indiana University, Bloomington, Indiana 47408, United States

Mycobacterium tuberculosis encodes 11 putative serine-threonine proteins Kinases (STPK) which regulates transcription, cell development and interaction with the host cells. From the 11 STPKs three kinases namely PknA, PknB and PknG have been related to the mycobacterial growth. From previous studies it has been observed that PknB is essential for mycobacterial growth and expressed during log phase of the growth and phosphorylates substrates involved in peptidoglycan biosynthesis. In recent years many high affinity inhibitors are reported for PknB. This paper describes how data fusion algorithms can identify top PknB inhibitors with high affinity. It has been observed that previous implementation of data fusion has shown effective enrichment of active compounds in both structure and ligand based approaches.

In this study we have used three types of data fusion ranking algorithms on the PknB dataset namely, sum rank, sum score and reciprocal rank. We have identified reciprocal rank algorithm is capable enough to select high affinity compounds earlier in a virtual screening process. Specifically, we found that the ranking of Pharmacophore search, ROCS and Glide XP fused with a reciprocal ranking algorithm not only outperforms structure and ligand based approaches but also capable of ranking actives better than the other two data fusion methods using the BEDROC, robust initial enhancement (RIE) and AUC metrics. We have also screened the Asinex database with best performing reciprocal rank algorithm to identify possible inhibitors for PknB. Using PCA we have shown that the predicted 45 compounds maps well to the PknB Inhibitor chemical space and can be further taken for experimental validation.

30 – 3D–QSAR using quantum-mechanics-based molecular interaction fields. Ahmed El Kerdawy (1), ahmed.elkerdawy@chemie.uni-erlangen.de, Stefan Güssregen (2), Hans Matter (2), Matthias Hennemann (1,3), Timothy Clark (1,3,4).

(1) Computer-Chemistry-Center, Friedrich-Alexander University Erlangen-Nürnberg, Erlangen, Bavaria 91052, Germany, (2) R&D, LGCR, Structure, Design and Informatics, Sanofi-Aventis Deutschland GmbH, Frankfurt am Main, Germany, (3) Interdisciplinary Center for Molecular Materials, Friedrich-Alexander University Erlangen-Nürnberg, Erlangen, Bavaria 91052, Germany, (4) Centre for Molecular Design, University of Portsmouth, Portsmouth, United Kingdom

The natural evolution of the different computer-aided drug design (CADD) methods involves a shift toward using quantum-mechanics (QM)-based approaches. This shift is not only the result of the ever growing computational power but also due to the need for more accurate and more informative approaches to describe molecular properties and binding characteristics than the currently available ones. QM approaches do not suffer from the limitations inherent to the ball-and-spring description and the fixed atom-centered charge approximation in the classical force fields mostly used by CADD methods. In this project we
introduce a protocol for shifting 3D-QSAR, one of the most widely used ligand-based drug design approaches, through using QM based molecular interaction fields (MIFs) which are the electron density (ρ), hydrogen bond donor field (HDF), hydrogen bond acceptor field (HAF) and molecular lipophilicity potential (MLP) to overcome the limitations of the current force-field-based MIFs. The average performance of the QM-MIFs (QMFA) models for nine data sets was found to be better than that of the conventional force-field-based MIFs models. In the individual data sets, the QMFA models always perform better than, or as well as, the conventional approaches. It is particularly encouraging that the relative performance of the QMFA models improves in the external validation.

62 – Standardized drug and pharmacological class network in ontology representation. Qian Zhu, zhu.qian@mayo.edu, Cui Tao, Christopher Chute.

Department of Health Sciences Research, Mayo Clinic, Rochester, MN 55901, United States

Dozens of drug terminologies and resources capture the drug and/or drug class information, ranging from their coverage and adequacy of representation. No transformative ways are available to link them together in a standard way, which hinders data integration and data representation for drug-related clinical and translational studies. In the meantime, ontology design and representation being widely used in the translational research offers more capabilities to discovery novel associations among entities underneath and consistency checkup. In this paper, we introduce a standardized drug and drug class network construction that integrates multiple drug terminological resources, using Anatomical Therapeutic Chemical (ATC) and National Drug File Reference Terminology (NDF-RT) as network backbone, and expanding with RxNorm and Structured Product Label (SPL). Meta-ontology will be established to capture drug, drug class and drug relevant information, in addition to the relationship among these entities. Instance layer will be built on the top of the meta-ontology.

64 – Teach-Discover-Treat: Round 2 competitions. Rommie E Amaro (1), ramaro@ucsd.edu, Johanna Jansen (2), Jane Tseng (4), Wendy Cornell (3).

(1) Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA 92093-0340, United States, (2) Novartis, United States, (3) Merck, United States, (4) National Taiwan University, Taiwan Republic of China

Teach-Discover-Treat (TDT) is an initiative to provide high quality computational chemistry tutorials that impact education and drug discovery for neglected diseases. The initiative is led by an international steering committee with members that work in academia and the pharmaceutical industry. TDT has 3 stated objectives: (1) Provide educational material in the form of online-accessible tutorials and presentations using neglected disease drug targets; (2) Provide access to developed models and materials that allow predictions of activity for small molecule compounds to be selected for chemical synthesis or biological assay;
and (3) Strengthen the community of scientists involved in open source drug discovery for neglected diseases and computational chemistry, through high-quality networking and outreach to the chemistry community. At the end of 2012, the initiative had produced a set of 6 tutorials through community action. Actual drug discovery activities around a Malaria challenge sponsored through TDT are ongoing. We will present a brief summary of our past accomplishments and describe the categories for our second round of competitions.

66 – Adventures in drug discovery: For now we see through a glass, darkly. Robert C Glen, rcg28@cam.ac.uk.

Department of Chemistry, University of Cambridge, Cambridge, Cambridgeshire CB21EW, United Kingdom

The fascinating thing about discovering drugs is that we learn more from our many mistakes than from our few successes – it is an incomplete puzzle with no clear answer. There are a few hundred successful drugs, and probably a few million attempts at drugs. The problem is that drug discovery is unfortunately multi-dimensional, with a response surface that is non-linear and disjoint, with polypharmacology and biological variability – and this is just for starters. Finding a new medicine is a really challenging problem. But there is hope – this is just the kind of problem we ought to make progress with using computational methods - embedding the transferrable nuggets of knowledge or technique in computer software not only helps what we do today, but encapsulates this knowledge for future generations of drug hunters. There are a multitude of methods available, but a few like CoMFA or Topomers stand out as innovative attempts at solving parts of the puzzle.

I would like to describe some problems, solutions and recent developments in computer-aided drug discovery applied to current projects at different stages of development.

83 – Teaching information literacy through an undergraduate laboratory project. Martin A Walker, walkerma@potsdam.edu.

Department of Chemistry, State University of New York at Potsdam, Potsdam, New York 13676, United States

One effective way to engage students with chemical information is through a lab project, designed such that their success depends on an effective and extensive search of the chemical literature. At the State University of New York at Potsdam, students in the introductory organic laboratory receive library training and perform a group search in a three hour session, as part of a larger project. This presentation will describe the successful outcomes, as well as the problems identified and lessons learned by this approach.

85 – Designing instruction activities to guide students through the research lifecycle: A science librarian approach Ye Li, liye@umich.edu.

Shapiro Science Library, University of Michigan, Ann Arbor, Michigan 48109, United States

Integrating research experiences into teaching and learning has gradually become essential in higher education. To support learning knowledge and skills needed for research, librarians traditionally provided instructions with a focus on information literacy to students. However, with the increasing importance of data and information in scientific research
in recent years, science librarians are expanding our scope of instruction to more areas throughout the research lifecycle. As information specialists with Science background, science librarians have a unique role in providing guidance and support to understand research process, communicate research ideas and results, obtain research funding, find, organize, evaluate and synthesize information as well as scientific data. In this study, we will first survey the literature to identify instructions which science librarians have designed to guide students through various steps of the research lifecycle. Then, two for-credit courses on research skills provided by the University of Michigan Library, one for first-/second-year undergraduates and another for post-second-year students in Science major, will be described and examined to demonstrate our successes and the challenges we encounter. In addition, we will also report other efforts dedicated to support students with research including managing references and data, presenting research results, writing scientific articles, and editing Wikipedia articles. Finally, we will map these instruction activities to the research lifecycle to illustrate our current strategy and identify possibilities for future development.

87 – Anything BUT overlooked: Librarians teaching scientific communication skills at the University of Florida. Donna T. Wrublewski (1), dtwrublewski@library.caltech.edu, Amy Buhler (2), Sara Gonzalez (2), Margeaux Johnson (2).

(1) California Institute of Technology, Pasadena, CA 91125, United States, (2) Marston Science Library, University of Florida, Gainesville, FL 32611, United States

Over the past 7 years, faculty science librarians at the University of Florida have developed and taught a three-credit Honors program course entitled “Discovering Research and Communicating Science”. The goal of this course is to prepare students (primarily freshmen) to begin undergraduate research, and thus teaches the ancillary skills often overlooked in advanced electives: searching and evaluating scientific literature, preparing a scientific poster, and writing scientific abstracts and papers. Guest researchers visit throughout the semester to expose students to undergraduate research opportunities and talk about success in research and other professional opportunities. This talk will discuss the motivation, organization, and ongoing development of the course over its different iterations. It will also present feedback from students in prior years, as well as potential relevance to standard chemical information literacy instruction.

92 – HPCC: A suitable solution for performing drug repositioning and preclinical pharmacological profiling Arnaud Sinan Karaboga (1), karaboga@harmonicpharma.com, Florent Petronin (1), Michel Souchet (1), Bernard Maigret (2).

(1) Department of Drug Repositioning, Harmonic Pharma, Villers lÂ¨s Nancy, France 54600, France, (2) orpailleur team, LORIA-CNRS UMR 7503, Vandoeuvre les Nancy, 54503, France

Here, we present a novel 3D molecular representation, namely Harmonic Pharma Chemistry Coefficient (HPCC), combining a ligand-centric pharmacophoric description projected onto a spherical harmonic based shape of a ligand [1]. First, we evaluate the performance of HPCC for molecular similarity assessment by discussing retrospective results obtained for some representative protein

93 – GES polypharmacology fingerprints: A novel and powerful drug repositioning tool Violeta I. Perez-Nueno (1), pereznueno@harmonicpharma.com, Arnaud S. Karaboga (1), Michel Souchet (1), Dave Ritchie (2).

(1) Harmonic Pharma, Villers les Nancy, France, (2) INRIA Nancy-Grand Est, Vandoeuvre les Nancy, France

We previously introduced the Gaussian Ensemble Screening (GES) approach to predict relationships between drug classes rapidly without requiring thousands of bootstrap comparisons as in current promiscuity prediction approaches [1]. Here, we present the GES polypharmacology fingerprint: the first fingerprint which codifies promiscuity information. It can be calculated for any ligand. Its length is variable and corresponds to the desired number of targets for which the promiscuity is investigated. The similarity between the 3D shapes and chemistry of ligands is measured with Parafit [2] and HPCC [3] and the promiscuity is quantified using GES. Hence, we obtain a consensus promiscuity representation based on the comparison of 3D shapes and chemistry of ligands. Here, we show as an example the GES polypharmacology fingerprint calculated for 800 targets linked to DrugBank [4] ligands. The performance of the approach is measured by comparing the present computational polypharmacology fingerprint with an in-house experimental polypharmacology fingerprint built using publicly available experimental data for the 800 targets that comprises the fingerprint. Matches between the computational and experimental polypharmacology fingerprints will be discussed.[ol][li]


96 WITHDRAWN

112 – Extraction, analysis, atom-mapping, classification, and naming of reactions from pharmaceutical ELNs. Roger Sayle, roger@nextmovesoftware.com, Daniel Lowe, Noel O’Boyle.
Electronic Laboratory Notebooks (ELNs) are widely used in the pharmaceutical industry for recording the details of chemical synthesis experiments. The primary use of this information is often for the capture of intellectual property for future patent filings, however this data can also be used in a number of additional applications, including synthetic accessibility calculations, reaction planning, and reaction yield prediction/optimization. Not only does a pharmaceutical ELN capture those classes of reactions suitable for small scale medicinal chemistry, but it is also uniquely a source of information on failed and poor yield reactions; an important class of data rarely found in the scientific literature or commercial reaction databases. This poster describes several of the technical chemoinformatics challenges in exploiting the wealth of synthetic chemistry information in ELNs. Starting with the hand-drawn sketches stored in relational databases, we describe the steps required to transform and normalize this data into a clean and annotated reaction database in an "open" file format such as MDL’s RD and RXN formats, or reaction SMILES. This process includes the tricky steps of reaction atom mapping, role assignment of reactants, reagents, catalysts and solvents, and the recognition of a reaction as an example of a known named reaction (Suzuki coupling, Diels-Alder cyclization, nitro reduction, chiral separation etc.) Novel (and improved) algorithms for each of these tasks will be described, and where appropriate compared to and benchmarked against previous methods and implementations.

117 – Chemotype approach to mapping the chemical landscape and exploring chemical–biological interactions within EPA’s ToxCast project. Rachelle J. Bienstock (1), bienstock.rachelle@epa.gov, Chihae Yang (2), Jim Rathman (2), Ann M Richard (3).

(1) IS&GS - Civil, Contractor Supporting the EPA National Center for Computational Toxicology (NCCT) Office of Research &Development US Environmental Protection Agency, Lockheed Martin, Research Triangle Park, NC 27711, United States, (2) Altamira LLC, Columbus, OH 43235, United States, (3) National Center for Computational Toxicology, Research Triangle Park, NC 27709, United States

U.S. EPA’s ToxCast project is employing high-throughput screening (HTS) technologies to profile thousands of chemicals that probe a wide diversity of biological targets, pathways and mechanisms related to toxicity. The current ToxCast chemical library is unprecedented in size (more than 1800 substances) and diversity, offering significant opportunities for cheminformatics contributions to toxicity modeling. However, the chemical diversity and nature of the HTS data sets present major challenges for QSAR modeling to contribute to the toxicity prediction problem. An approach employing a standard set of toxicity-informed feature sets, or chemotypes, is being developed to resolve the global chemical landscape into groupings of potential biological relevance. Use of these chemotypes in conjunction with biological knowledge and adverse-outcome pathway hypotheses, offers a means to focus and constrain modeling efforts into potentially productive areas of chemical and biological space, thereby improving modeling success and interpretability. Abstract does not represent EPA policy.
Tuesday, September 10, 2013

Herman Skolnik Award Symposium – AM Session

Indiana Convention Center
Room: 140

Richard Cramer, Organizers
Brian Masek, Presiding

8:30 am – 11:50 am

8:30 – Introductory Remarks

8:35 – 65 – Adventures in CoMFAland
Robert D Clark, bob@simulations-plus.com

Department of Life Sciences, Simulations Plus, Inc., Lancaster, CA 93534, United States

My introduction to Comparative Molecular Field Analysis (CoMFA) and Dick Cramer came in the early 1990’s while I was still working at Monsanto Agricultural Co. I continued working with the technology and the man for most of the next 20 years, many of them spent at Tripos Inc. This presentation will describe the history of both from the perspective of someone directly involved in the many and varied developments along the way.

9:05 – 66 – Adventures in drug discovery: For now we see through a glass, darkly.
Robert C Glen, rcg28@cam.ac.uk

Department of Chemistry, University of Cambridge, Cambridge, Cambridgeshire CB21EW, United Kingdom

The fascinating thing about discovering drugs is that we learn more from our many mistakes than from our few successes – it is an incomplete puzzle with no clear answer. There are a few hundred successful drugs, and probably a few million attempts at drugs. The problem is that drug discovery is unfortunately multi-dimensional, with a response surface that is non-linear and disjoint, with polypharmacology and biological variability – and this is just for starters. Finding a new medicine is a really challenging problem. But there is hope – this is just the kind of problem we ought to make progress with using computational methods - embedding the transferrable nuggets of knowledge or technique in computer software not only helps what we do today, but encapsulates this knowledge for future generations of drug hunters. There are a multitude of methods available, but a few like CoMFA or Topomers stand out as innovative attempts at solving parts of the puzzle. I would like to describe some problems, solutions and recent developments in computer-aided drug discovery applied to current projects at different stages of development.

Yvonne C Martin, yvonnecmartin@comcast.net

Martin Consulting, Waukegan, Illinois 60087, United States

In 1963 Hansch and Fujita invented 2D QSAR, which signaled the conversion of drug design from chemists’ intuition and ease of synthesis to property-based considerations aided by computer analysis. Left as by-standers were researchers who ignored the use of statistical analysis with computers or who ignored the importance of hydrophobicity to drug potency. Some twenty-five years later Richard Cramer III and colleagues united molecular modelling
and QSAR with the CoMFA method that showed that it is possible to correlate the potency of existing and forecast the potency of untested compounds with their 3D properties. Left as by-standers were researchers who focussed on complex descriptions of 3D properties or who ignored the use of multivariate statistical methods. Five years after that, Martin and colleagues showed that it is possible to use a computer to identify the 3D pharmacophores present in a set of diverse molecules, thus providing a starting hypothesis for CoMFA analysis. Left as by-standers were those who developed methods that required the user to specify the corresponding atoms in the molecules or who focussed on the 3D structures of the molecules, not their properties in 3D space. In each of these examples, the paradigm shift was catalyzed by the focus of the investigators to solve a problem.

10:05 – Intermission


Department of Pharmecutical Sciences, The University of New Mexico, Albuquerque, NM 87131-0001, United States

Dick Cramer and I began to discuss the future of QSAR analysis in the mid- to late- 1970s. There was little we did agree upon, but we were in concurrence that 3D information about molecules somehow needed to be included in the then 2D-QSAR paradigm. Much of our discussion and focus was directed at issues that remain familiar even today - a) how to represent the 3D information as descriptors, b) how to select ‘active’ conformations, c) what is the best way to select alignments, and, d) is there a good approach to develop a scheme to simulatanuously do data reduction and model/function optimization. Obviously Dick was enormously successful with CoMFA by cleverly bringing together diverse computational and statistical methods to first generate and then parse through and neatly package large amounts of 3D molecular field data to generate a significant 3D-QSAR. At the same time, I struggled with first developing a methodology called molecular shape analysis, MSA, and making it operational. Subsequently, MSA morphed into what is now called 4D-QSAR analysis which can be considered a hybrid of CoMFA and MSA. Along the way, namely the last 35 years or so, the interesting discussions between Dick and I hopefully, on occassion, have served as substrates leading to advances, ideas and concepts, as well as more interesting discussions among our colleagues. I don’t want this interplay between Dick and I to end so I’m going to mention in this presentation some recent advances, or perhaps better stated, hidden features, of 4D-QSAR analysis relating to alignment, pharmacophore delineation, ‘active conformation’ and mixing of different types and classes of descriptors within a 3D-QSAR formalism. Hopefully, this will once again garner interest from Dick for more discussion.

10:50 – 69 – Evolution of QSAR from regression analysis to physical modeling. Ajay N Jain, ajain@jainlab.org.

Department of Bioengineering and Therapeutic Sciences, UCSF, San Francisco, CA 94158, United States

This symposium honors a scientist whose work began a shift from predictive modeling of chemical and biological properties of molecules based on...
correlative analyses to modeling of biological activity in a manner related to underlying physical principles. Our work has been greatly influenced by the trail that was initially broken by CoMFA. Over the past twenty years, we have followed a course of increasingly more physically realistic model induction, beginning with abstract machine-learning models that addressed ligand pose variation and continuing now with methods that combine protein structural information with ligand activity information to produce truly physical models of binding sites. Such models are capable of accurate binding affinity predictions and accurate predictions of bioactive molecular poses.


Department of Biological Sciences, University of Delaware, Newark, DE 19716, United States

Science may be understood as a method of organized analysis which follows principles of objectivity, reproducibility and testability of clearly defined hypotheses. There is no limitation to specific topics such as Chemistry, Physics, or Biology. One non-traditional area in which the scientific approach has had great success is the analysis of baseball performance. Terms such as “baseball analytics” and “sabermetrics” have become increasingly common as science has moved into the sports world. There are three especially interesting aspects of scientific baseball analysis: 1) The innovations and sophisticated thought originated outside of the professional teams, many of which have been slow to accept what they see as intrusions; 2) Empirical observations have been important, but there has been a substantial component of modeling as well; 3) The collection of high quality, reliable data has been an essential underpinning to the entire effort. Retrosheet is a volunteer organization which has an extensive database of detailed baseball data that has been used in many studies. It is therefore not at all surprising that professional scientists and mathematicians have combined a passion for baseball with this rigorous analysis of a game. As a result, most teams now use a scientific approach to some degree, with the book (and later movie) “Money Ball” as a clear example along with the aggressive and successful use of these techniques by the Boston Red Sox. But the value of science in baseball is much greater, as a deeper and expanded understanding of the game’s complexities has enhanced appreciation and enjoyment for fans and professionals as well as analysts.

Tuesday, September 10, 2013

Herman Skolnik Award Symposium – PM Session

Indiana Convention Center
Room: 140

Richard Cramer, Organizers
Terry Stouch, Presiding

2:00 pm – 4:45 pm

2:00 – Introductory Remarks


Department of Chemistry and Biochemistry, University of California
Santa Cruz, Santa Cruz, CA 95064, United States
This paper explores the technology of computer-assisted synthesis planning, its beginning, evolution, and impact. Once chemists could communicate in their natural language (structural diagrams) and the computer could carry out the symbolic algebra of chemical reactions we entered a new era for organic chemists and chemical information processing. Chemists had a new co-worker, one that would work faster, smarter, and cheaper each year. Fast forward 44 years, has the new partnership achieved all it was capable of? What can we expect for the future?

2:35 – 72 – Think local, act global: Some challenges in cheminformatics and drug research. Tudor I Oprea, toprea@salud.unm.edu.

Department of Internal Medicine, Translational Informatics Division, University of New Mexico School of Medicine, Albuquerque, NM 87131, United States


3:05 – Intermission

3:20 – 73 – From library design to off-target prediction: A wide array of topomer applications. Bernd Wendt, bernd.wendt@certara.com.

Certara, Munich, Deutschland 81829, Germany

The topomer is a molecular descriptor that provides one solution to the molecular alignment problem. It produces a highly consistent set of 3D representations of fragments and allows for 3D alignment-based comparisons of molecules. The topomer has been widely used in various discovery research applications such as library design, virtual screening, 3D-QSARand off-target prediction. Examples from several applications will be illustrated and discussed1,2. [ol][li]Wendt B, Mülbaierr M, Wawro S, Schultes C, Alonso J, Janssen B, Lewis J (2011) J Med Chem 54:3982-3986 Wendt B, Uhrig U, B F (2011) J Chem Inf Model 51:843-851


Tripos, Certara, Santa Fe, NM 87507, United States

3D-QSAR’s fundamental challenge is generating appropriate 3D superpositions, or “alignments”, of training and test set candidate structures. Whole Template CoMFA alignment is based on one or more template structures, whose conformations may be experimentally determined and/or pharmacophorically hypothesized. Alignment of a candidate structure first identifies and overlays the candidate bond having maximal “similarity” to any bond in any template structure, then copies the coordinates from all the “matching” atoms within that template to the corresponding candidate atoms, and finally positions the remaining candidate atoms by topomer canonicalization. Virtues of this new
protocol include its full utilization of both structural and SAR information, ready interpretability and applicability, objectivity, and (foreseeably complete) automatibility. In addition to this method and some sample applications, its potential relevance to such fundamental QSAR challenges as the scope and reliability of a particular QSAR’s predictions will be discussed.

4:35 – Award Presentation

Wednesday, September 11, 2013

Exchangeable Molecular and Analytical Data Formats and their Importance in Facilitating Data Exchange – AM Session Exchanging Molecular Data

Indiana Convention Center
Room: 140

Cosponsored by COMP

Antony Williams, Robert Lancashire, Organizers
Antony Williams, Presiding

8:05 am – 12:00 pm

8:05 – Introductory Remarks

8:10 – 75 – Cheminformatics runs on molfiles and its siblings: There is a molfile for that. Keith T Taylor, keith.taylor@accelrys.com.

Accelrys Inc, San Ramon, CA 94583, United States

The molfile is ubiquitous in cheminformatics; it is virtually mandatory that an application can read and write them usually concatenated in the form of a SDfile. The molfile belongs to the class of representation known as a connection table in contrast to line notations such as SMILES and InChI. The origins, and evolution of the molfile, and what benefits does it deliver to the user experience will be discussed.

8:40 – 76 – Reading and writing molecular file formats for data exchange of small molecules, biopolymers, and reactions Roger A Sayle, roger@nextmovesoftware.com.

NextMove Software, Cambridge, CAMBS CB4 0EY, United Kingdom

Modern pharmaceutical companies have hundreds of information systems and applications for keeping track and processing of compounds. From registration systems, to electronic lab notebooks, to inventory management, to predictive chemistry applications such as ADMET QSAR modeling and virtual screening. The plumbing that holds these myriad systems together is the interchange of files encoding the connection table representations of molecules. Over the years a number of de facto standard file formats have gained, if not popularity, widespread usage between systems from multiple vendors. These include MDL’s mol, SD, rxn and RD file formats, Daylight’s SMILES strings, Tripos’ Mol2 files, PDB files, ChemDraw and ISIS/Draw sketches etc. Curiously, the more widely used a molecular file format, the more misunderstood and maligned it becomes. Alas often poor reputations are not the fault of the file formats themselves (or their designers) but the inherent complexity of representing molecules in a computer, and ignorance of how to correctly interpret their content. If a file format were not fit for some purpose, it would not have become popular and would likely have been be displaced by a better competing representation. In this talk, I
describe features of various popular file formats that solve important challenges in exchanging chemical information, but alas are often poorly or not widely implemented. Hopefully, raising awareness of such features may help them become more widely implemented and adopted. As one example, I shall present details of community-wide efforts to improve the handling of implicit hydrogen valence in MDL mol files. An experiment to compare the interpretation of MDL V2000 files across more than twenty cheminformatics tools revealed a significant number of differences in interpretation. As a result, many issues uncovered by this investigation have since been fixed by vendors and software developers, including RDKit, OpenBabel, Optibrium, CACTVS, NAOMI and MayaChemTools.

9:10 – Intermission

9:20 – 77 – Facilitating accurate chemical data interconversion using Open Babel: The good, the bad, and the painful. Geoffrey R Hutchison, geoffh@pitt.edu.

Department of Chemistry, University of Pittsburgh, Pittsburgh, PA 15260, United States

The Open Babel project is a widely-used open source chemical toolbox designed to speak the many languages of chemical data. It’s an open, collaborative project allowing anyone to search, convert, analyze, or store data from molecular modeling, chemistry, solid-state materials, biochemistry, or related areas. In the latest version, it supports over 100 file formats for molecular and chemical data with wide ranges of chemical representations. We will discuss what works, what doesn’t, and ways to move forward with accurate data interchange.

9:50 – 78 – Exchanging chemical structures and ELN data: CDX, CDXML, and other formats. David Gosalvez, Alex Jewett, Phil McHale, phil.mchale@perkinelmer.com, Churl Oh, Rudy Potenzone, rudy.potenzone@perkinelmer.com, Chris Strassel.

Informatics, PerkinElmer Inc., Waltham, MA 02451, United States

As research becomes more distributed, there is a pressing need to be able to share and exchange data in a robust way with no loss in information or fidelity. In particular, dispersed research groups and organizations with outsourced synthesis or testing partners need ways to exchange chemical structures, synthetic pathways and assay results, many of which will have been captured and stored in electronic lab notebooks (ELNs). This paper will describe the open cdx and cdxml formats that can be used to share molecular information, including not just connection tables but also orientation, layout, fonts and non-structural elements to retain complete fidelity with the original drawing. We will also describe emerging XML-based standards for secure and accurate exchange of experimental data between ELNs.

10:20 – 79 – InChI: Recent developments in the worldwide chemical structure identifier standard. Stephen Heller, steve@hellers.com.

InChI Trust, Silver Spring, MD 20902, United States

The IUPAC InChI/InChIKey project has evolved to the point that over 125 million InChIs and InChIKeys are now in databases (searchable over the web via Google, Blekko, and other search engines), such as ChemSpider, Reaxys, NIH/NCI, and NIH/NLM/PubChem. There
are now more InChIs and InChIKeys searchable and available on the Internet than any other chemical structure representation. The InChI Trust, an independent UK not-for-profit entity supported and paid for by the chemical information and publishing community and those who use and benefit from the InChI algorithm has been funding the ongoing programming of this effort. The mission of the Trust is quite simple and limited; its sole purpose is to create and support administratively and financially a scientifically robust and comprehensive InChI algorithm and related standards and protocols.

This presentation will describe the current technical state of the InChI algorithm including InChIs for reactions (RInChI), Markush InChI, InChIs for polymers, the InChI QR app, InChI education videos, and the InChI Certification Suite software which that InChIs are properly and consistently generated throughout the world.

10:50 – Intermission

11:00 – 80 – Data exchange caveats and particulars, the devil is in the details. 
Evan Bolton, bolton@ncbi.nlm.nih.gov.

National Center for Biotechnology Information, National Institutes of Health, Bethesda, MD 20894, United States

Data exchange is a key aspect of building community driven resources, such as those for chemical biology information. As an example, PubChem has more than 200 contributors providing over 110 million substance descriptions, 45 million unique chemicals, and over 200 million biological testing results from more than 200 contributors. Data exchange standards help to facilitate effective communication of information. Using chemical structure data exchange as an example, caveats and pitfalls involved in the use of various format flavors and data representation approaches will be discussed.

11:30 – 81 – Pistoia Alliance and the emerging HELM standard at the “dawn of the ADC informatics era”. Sergio Rotsein (1,2), John Wise (1,3), Claire Ballamy (1), Michael Braxenthaler (1,4), Barry Bunin (1,5), bbunin@collaboratedrug.com.

(1) Pistoia Alliance, Pistoia, Pistoia, Italy, (2) Department of Research Business Technology, Pfizer, Cambridge, MA 02140, United States, (3) Pharma Logistics Ltd., Mundelein, IL 60060, United States, (4) Pharma Research and Early Development Informatics, Roche, Nutley, NJ 07110, United States, (5) Department of Corporate Development, Collaborative Drug Discovery (CDD), Inc., Burlingame, CA 94010, United States

The recent increase in utilization of complex biomolecules such as Antibody-Drug Conjugates (ADCs) as therapeutic agents has revealed a substantial gap in the biopharmaceutical industry’s portfolio of informatics tools and methods, most of which were designed to work primarily with either small molecules or unmodified and unconjugated amino acid and nucleotide sequences. The Hierarchical Editing Language for Macromolecules (HELM), along with a software toolkit that leverages that language was developed by Pfizer researchers to address this gap (see J. Chem. Inf. Model 2012, 52, 2796-2806). HELM and related technologies such as SCSR (Self-Contained Sequence Representation, see: J Chem Inf Model. 2011 51, 2186-208.), enable the representation of unnatural, conjugated or otherwise modified building blocks in biopolymers such as oligonucleotides,
peptides, proteins and ADCs. HELM is being released into the Open Source and the Pistoia Alliance has initiated a project that has the goals of (1) facilitating this release, (2) adopting HELM as an industry standard for the manipulation and exchange of complex biomolecule data, and (3) setting up the organizational infrastructure to govern the future development of the standard. Given the recent FDA approval of the Antibody Drug Conjugate Kadcyla (Ado-Trastuzumab Emtansine) for Her2-Positive Metastatic Breast Cancer, one can reasonably assume that standards like HELM and associated software tools will play an increasingly prominent role as part of the modern drug discovery informatics arsenal.

Wednesday, September 11, 2013

Before and After Lab: Instructing Students in 'Non–Chemical' Research Skills – AM Session

Indiana Convention Center
Room: 141

Cosponsored by CHED, YCC
Andrea Twiss–Brooks, Charles Huber, Organizers
Charles Huber, Presiding

8:10 am – 12:00 pm

8:10 – Introductory Remarks

8:15 – 82 – Chemical information across San Diego County: A community college and university library collaboration for an independent synthesis project. Teri M Vogel (1), tmvogel@ucsd.edu, Cynthia B Gilley (2).

(1) Library, University of California San Diego, La Jolla, CA 92093, United States
(2) Department of Chemistry, Palomar Community College, San Marcos, CA 92069, United States

Community college students have a number of freely available chemical information resources at their disposal. However, for an advanced assignment requiring greater access to the chemical literature, they may find themselves at a disadvantage compared with students at nearby institutions whose libraries provide more licensed electronic resources. In February 2013, a Palomar chemistry instructor partnered with the UCSD chemistry librarian to connect her second semester organic chemistry class with the resources and assistance needed for an independent synthesis project, offering the students a richer, more real-life experience, as well as an introduction to the library and the chemical literature. In this presentation, we will share how this collaboration came about, how we prepared for and taught the library instruction session, student outcomes and feedback, and what we have learned for future classes.

8:45 – 83 – Teaching information literacy through an undergraduate laboratory project. Martin A Walker, walkera@potsdam.edu.

Department of Chemistry, State University of New York at Potsdam, Potsdam, New York 13676, United States

One effective way to engage students with chemical information is through a lab project, designed such that their success depends on an effective and extensive search of the chemical literature. At the State University of New York at Potsdam, students in the introductory organic laboratory receive library training and perform a group search in a three hour
session, as part of a larger project. This presentation will describe the successful outcomes, as well as the problems identified and lessons learned by this approach.

9:15 – 84 – Integrating citations as a teaching element into chemistry information literacy training sessions. Shu Guo, guo1s@cmich.edu.

Reference Department, Central Michigan University, Mt. Pleasant, MI 48859, United States

Since science/chemistry librarians generally do not get too much teaching time in regular chemistry classrooms, we very like will focus our teaching on how to search for chemical compounds and how to obtain information on physical, chemical and other properties of the chemical compounds. The information students obtained are still in pieces, although they may know how to search for reference books and how to search several databases to locate journals and articles. At least these are the basics science/chemistry librarians have been emphasizing during our regular chemistry information literacy training sessions. At Central Michigan University, four chemistry information literacy training sessions were integrated into Chem 349, an Organic Chemistry lab course, a core course required for all students majored in Chemistry and Biochemistry and for some students majored in Biomedical Sciences and Biology for certain concentrations. Through the four sessions, students get the training on different searching techniques including chemical structure searching and they also learn how to search databases: SciFinder, Reaxys and Web of Science. Most importantly, Science Librarian has integrated the citation as a teaching element into the Chemistry information training sessions. Students have learned: how to read a research article and locate the source citations listed in the reference section; how to interpret a citation to determine if it is a book, a chapter in an edited book, a report, a patent or a journal article; how to use library catalog and journal citation linker to locate a specific item in the library collection; how to use the cited and citing reference lists to locate related sources on a research topic. Students feel more confident after finishing the lectures and assignments associated with each lectures, and they can apply the knowledge just obtained to “real lab” problems right away.

9:45 – Intermission

10:00 – 85 – Designing instruction activities to guide students through the research lifecycle: A science librarian approach. Ye Li, liye@umich.edu.

Shapiro Science Library, University of Michigan, Ann Arbor, Michigan 48109, United States

Integrating research experiences into teaching and learning has gradually become essential in higher education. To support learning knowledge and skills needed for research, librarians traditionally provided instructions with a focus on information literacy to students. However, with the increasing importance of data and information in scientific research in recent years, science librarians are expanding our scope of instruction to more areas throughout the research lifecycle. As information specialists with Science background, science librarians have a unique role in providing guidance and support to understand research process, communicate research ideas and results, obtain research funding, find, organize,
evaluate and synthesize information as well as scientific data. In this study, we will first survey the literature to identify instructions which science librarians have designed to guide students through various steps of the research lifecycle. Then, two for-credit courses on research skills provided by the University of Michigan Library, one for first-/second-year undergraduates and another for post-second-year students in Science major, will be described and examined to demonstrate our successes and the challenges we encounter. In addition, we will also report other efforts dedicated to support students with research including managing references and data, presenting research results, writing scientific articles, and editing Wikipedia articles. Finally, we will map these instruction activities to the research lifecycle to illustrate our current strategy and identify possibilities for future development.

10:30 – 86 – “I can just copy this, right?”: Introducing students to copyright. Charles F. Huber, huber@library.ucsb.edu.

As prospective users and creators of information, students need to learn about copyright. Among the aspects of copyright relevant to today’s students are: what is copyrightable?, copyright and fair use; author’s rights and publication; and what “open access” means. Useful resources for teaching about copyright will be described.

11:00 – 87 – Anything BUT overlooked: Librarians teaching scientific communication skills at the University of Florida. Donna T. Wrublewsiki, (1) dtwrublewski@library.caltech.edu, Amy Buhler (2), Sara Gonzalez (2), Margeaux Johnson (2).

(1) California Institute of Technology, Pasadena, CA 91125, United States, (2) Marston Science Library, University of Florida, Gainesville, FL 32611, United States

Over the past 7 years, faculty science librarians at the University of Florida have developed and taught a three-credit Honors program course entitled “Discovering Research and Communicating Science”. The goal of this course is to prepare students (primarily freshmen) to begin undergraduate research, and thus teaches the ancillary skills often overlooked in advanced electives: searching and evaluating scientific literature, preparing a scientific poster, and writing scientific abstracts and papers. Guest researchers visit throughout the semester to expose students to undergraduate research opportunities and talk about success in research and other professional opportunities. This talk will discuss the motivation, organization, and ongoing development of the course over its different iterations. It will also present feedback from students in prior years, as well as potential relevance to standard chemical information literacy instruction.

11:30 – 88 – Introducing Electronic Laboratory Notebooks (ELNs) to students and researchers at the University of Maryland College Park.

Svetla Baykoucheva (1), sbaykouc@umd.edu, Lee Friedman (2).

(1) White Memorial Chemistry Library, University of Maryland College Park, College Park, Maryland 20472, United States, (2) Department of Chemistry and Biochemistry, University of Maryland College Park, College Park, Maryland 20472, United States
Electronic Laboratory Notebooks (ELNs) show a lot of promise, and it is not surprising that more and more companies are embracing this technology. ELNs can make projects more organized, streamlined, easier to search and share, and ultimately, save a lot of time and make research better. The University of Maryland Libraries in College Park started introducing ELNs to students and faculty in 2011. Classes were offered to all students and faculty, and different models of ELNs were demonstrated. Parts of library instruction classes taught in one undergraduate and one graduate course were devoted to the use and benefits of ELNs. In the 2013 Spring semester, a chemistry librarian partnered with a course instructor to develop a pilot project for an instrumental undergraduate chemistry course. For this project, a LabArchives ELN (Classroom Edition) was used as a model. All experimental protocols and reports from a hard-copy course handbook were digitized and uploaded to an ELN course notebook. Students had to use the ELN to access experiment protocols; complete and submit lab reports; open, upload, edit, and share Word, Excel, PDF files; and demonstrate that they had mastered the basic features of the ELN. This presentation will discuss the results from the pilot project and outline possible future strategies for introducing this advanced technology more broadly to chemistry majors and graduate courses.

Wednesday, September 11, 2013

Computational Profiling and Repositioning as Promising New Ways of Drug Development – AM Session

Indiana Convention Center
Room: 142
Selectivity is an important issue in drug development. To better understand the reasons why bioactive compounds bind to different proteins, we identified in the Protein Databank 247 “drug-like” ligands in complex with two or more distinct protein targets. Studying the similarity between the ligand-binding sites in the different targets revealed that the lack of selectivity of a ligand can be due (i) to the fact that Nature has created the same binding pocket in different proteins which do not necessarily have otherwise sequence or fold similarity, or (ii) to specific characteristics of the ligand itself. In particular, we demonstrated that ligands can adapt to different protein environments by changing their conformation, or by using different chemical moieties to anchor to different targets, or by adopting unusual binding modes. Lastly, we suggested possible relationships between structure and promiscuity.

The above figure gives an example of promiscuous ligand: the three-dimensional structure of the inhibitor sb220025 in complex with two different map kinases, p38 (PDB: 1bl7) and erk2 (PDB: 3erk).

8:55 – 91 – Predicting and testing drug off- and on-targets. Brian Shoichet, bshoichet@gmail.com.

Faculty of Pharmacy and Ontario Institute for Cancer Research, University of Toronto, Toronto, Ontario M5S 3E1, Canada

Pharmacological targets share ligand similarities that are not reflected by evolutionary relationships, and drugs and reagents will often modulate targets that are surprising bioinformatically. Many of these unexpected cross-pharmacologies may be predicted chemoinformatically. Using one such method, we predict and experimentally test both mechanism of action targets—for those drugs and tool compounds for which these are not known, or known wrong—and also targets that contribute to the adverse reactions of drugs. As successful as this chemoinformatics program of research has been, a crucial weakness is its reliance on inferences drawn from known ligand-target associations. We are therefore exploring using virtual lists of ligands to define receptor ligand recognition patterns, comparing receptors based on docking hit-lists, and testing the resulting associations experimentally. Whereas this approach is freighted with the problems of molecular docking, recent results suggest that it might not be completely insane. We are grateful for collaborations with the groups of Prof Brian Roth (UNC-Chapel Hill) and Dr. Laszlo Urban (Novartis, Cambridge), without whom the projects would have been impossible. B. Shoichet is a founder of SeaChange Pharmaceuticals and declares a potential financial conflict of interest.

9:20 – Intermission

Here, we present a novel 3D molecular representation, namely Harmonic Pharma Chemistry Coefficient (HPCC), combining a ligand-centric pharmacophoric description projected onto a spherical harmonic based shape of a ligand [1]. First, we evaluate the performance of HPCC for molecular similarity assessment by discussing retrospective results obtained for some representative protein targets from the commonly used and publicly available Directory of Useful Decoys (DUD) data set comprising over 100,000 compounds distributed across 40 protein targets of therapeutic interest [2]. Second, we show the efficiency of HPCC in prospective conditions with case studies where HPCC was successfully applied to repurpose drugs and preclinical compounds. 1. Benchmarking of HPCC: A novel 3D molecular representation combining shape and pharmacophoric descriptors for efficient molecular similarity assessments. Karaboga, A.S.; Petronin, F.; Marchetti, G.; Souchet, M.; Maigret, B. (2013) J. Mol. Graph. Model. 41, 20-30. 2. Benchmarking sets for molecular docking. Huang, N.; Shoichet, B.K.; Irwin, J.J. (2006) J. Med. Chem. 49, 6789-6801.

We previously introduced the Gaussian Ensemble Screening (GES) approach to predict relationships between drug classes rapidly without requiring thousands of bootstrap comparisons as in current promiscuity prediction approaches [1]. Here, we present the GES polypharmacology fingerprint: the first fingerprint which codifies promiscuity information. It can be calculated for any ligand. Its length is variable and corresponds to the desired number of targets for which the promiscuity is investigated. The similarity between the 3D shapes and chemistry of ligands is measured with Parafit [2] and HPCC [3] and the promiscuity is quantified using GES. Hence, we obtain a consensus promiscuity representation based on the comparison of 3D shapes and chemistry of ligands. Here, we show as an example the GES polypharmacology fingerprint calculated for 800 targets linked to DrugBank [4] ligands. The performance of the approach is measured by comparing the present computational polypharmacology fingerprint with an in-house experimental polypharmacology fingerprint built using publicly available experimental data for the 800 targets that comprises the fingerprint. Matches between the computational and experimental polypharmacology fingerprints will be discussed.


Developmental & Molecular Pathways, Novartis Institutes for BioMedical Research, Cambridge, MA 01239, United States

Computational prediction of compound targets is a growing discipline that takes advantage of QSAR approaches applied to large-scale global pharmacology data. For several years, multiple-category Bayesian models have been a mainstay in target prediction due to their ease of computation and application. Bayesian models trained on targets, domain, and other categories are highlighted, including successful deconvolution of targets for compounds discovered in phenotypic screens. Further application of Bayes models as virtual affinity fingerprints in similarity searching is exemplified.

10:45 – Intermission

10:55 – Effect and target profile prediction by interaction pattern based drug design. Zoltan Simon (1,2), Agnes Peragovics (3), Laszlo Vegner (3), Balazs Jelinek (1,4), Peter Hari (1,2), Istvan Bitter (1,5), Pal Czobor (1,5), Andras Malnasi-Csizmadia, malnalab@yahoo.com.

(1) Drugmotif Ltd., Veresegyhaz, Hungary, (2) Printnet Ltd., Budapest, Hungary, (3) Department of Biochemistry, Eotvos Lorand University, Budapest, Hungary, (4) Molecular Biophysics Research Group, Hungarian Academy of Sciences - Eotvos Lorand University, Budapest, Hungary, (5) Department of Psychiatry and Psychotherapy, Semmelweis University, Budapest, Hungary

Our Drug Profile Matching (DPM) approach relates complex drug-protein interaction profiles with effect and target profiles of drugs and druglike molecules. DPM is based on the docking profiles of cca. 1,200 FDA-approved small-molecule drugs against a set of non-target proteins and creates bioactivity predictions based on this pattern. The effectiveness of this approach for the prediction of 129 therapeutic effect categories and 77 targets was measured by probability values calculated with linear discriminant analysis and validated by 10-fold cross-validation. The average AUC values of the validations for all predictions of the therapeutic effects and targets were 0.791±0.147 and 0.839±0.081, respectively. These results proved the applicability of DPM for drug repositioning even in effect categories containing a structurally diverse set of drugs. In the case of target predictions, 79% of the known drug-target interactions were correctly predicted by DPM, and additionally 1,074 new drug-target interactions were suggested. Based on the experimental testing of ACE and COX inhibitory effects and dopaminergic activities, the positive hit rates for the newly predicted molecules were between 47-84%. Currently we are testing these activities on a set of 600 000 druglike compounds.

Drug repositioning holds a tremendous potential to cost-effectively explore drugs with favorable pharmacokinetics and safety profiles for emerging targets. Here we utilized novel integrative modeling approaches to identify new therapeutic indications for several FDA-approved drugs against an unexploited colorectal cancer target, TNIK. We first collected a dataset of 102 compounds with known TNIK activities and performed molecular docking against TNIK. Based on the alignment of their docked poses, we derived QSAR models using CoMSIA and kNN methods. Using these models, we identified a set of drugs (e.g., sunitinib) which strongly inhibit TNIK activities upon screening of 1,448 FDA-approved small-molecule drugs. These predictions were further validated using the KINOMEscan profiling platform, confirming that their binding affinities (e.g., Kd) to TNIK are in the range of $10\mu\text{M}$ to $50\text{nM}$. The subsequent modeling analyses proved that these agents acquire favorable molecular characteristics to inhibit TNIK kinase activities.

Solute Carrier (SLC) transporters are membrane proteins that control the uptake and efflux of a broad spectrum of substrates, such as nutrients, toxins, and prescription drugs. In human, there are 386 SLC transporters that can be drug targets themselves or be responsible for absorption, disposition, metabolism, and excretion of drugs. We first perform a comprehensive comparison of the SLC transporters to inform attempts to model their atomic structures, a prerequisite for structure-based ligand discovery. We then describe an integrated computational and experimental approach for identifying transporter-small molecule interactions. Particularly, we use comparative modeling and virtual screening, followed by experimental validation by measuring uptake kinetics, to identify interactions between SLC transporters and small molecules ligands, including prescription drugs, metabolites, and fragment-like compounds. For example, we discovered that several existing prescription drugs interact with the norepinephrine transporter, NET, which may explain some of the pharmacological effects (i.e., efficacy and/or side effects) of these drugs via polypharmacology. Our combined theoretical and experimental approach is generally applicable to structural characterization of protein families other than transporters, including receptors, ion-channels, and enzymes, as well as their interactions with small molecule ligands.
Room: 140

Cosponsored by COMP

Antony Williams, Robert Lancashire, Organizers
Robert Lancashire, Presiding

1:30 pm – 4:15 pm


(1) Research, Development & Innovation, AkzoNobel Chemicals bv Zutphenseweg 10, Deventer, The Netherlands, (2) Department of Chemistry, The University of the West Indies, Kingston, St Andrew Kgn 7, Jamaica

The Joint Committee on Atomic and Molecular Physical Data started as a Task Force on Spectral Data Portability under the direction of Paul A. Wilks, Jr., at the Pittsburgh Conference (Pitcon) of 1983. The scope of JCAMP was originally as follows: “The Joint Committee will generate, collect, evaluate, edit, and approve the publication and encourage the distribution of atomic and molecular physical data in suitable form to serve as references for pure compounds and mixtures”. The first objective of the Task Force was to design a standard file format for exchange of infrared spectra between vendor data systems that used different proprietary file formats. Data exchange capability was in demand by end-users who wished to transfer spectra between different spectrometers in their own and other laboratories. In 1988, the first JCAMP-DX spectroscopic data format was published (for IR). Since then the protocols for a range of techniques have been developed and published, the latest being for circular dichroism (2012) A review of the standards and prospects for the future will be highlighted

2:00 – 99 – Knowledge sharing or what I learned in first grade. Michael Boruta, michael.boruta@acdlabs.com.

Advanced Chemistry Development, Inc., Toronto, Ontario M5C 1B5, Canada

Sharing is a fundamental part of our early development in part because it helps maintain civility, but primarily because it broadens our experiences. In the scientific world sharing information is one of the key factors in the advancement of science. The ability the share or exchange information with colleagues improves our productivity, reduces our costs and enables us to gain insights into a variety of problems. This talk will look at a few examples where the ability to share information has a direct impact on our lives as scientists.

2:30 – 100 – Long wait for exchangeable data formats vs. the evolution of data. Clemens Anklin, Clemens.Anklin@bruker-biospin.com.

Bruker Biospin Corp., Billerica, MA 01821, United States

Instrument manufacturers typically provide a proprietary format for the data produced on their equipment. It is up to the individual company to decide on the amount of release of information about the format. Exchangeable data formats are often welcomed as they provide easy access and portability. They can eliminate the need for individual data import and export functions.

The main problems with most proposed data formats has been the speed at which these develop and that rate at which
they spread through the community. Unfortunately they tend to lag behind instrument and methods development. JCAMP can serve as a typical example. The current JCAMP5.0 standard does not support anything beyond 1D NMR, where NMR itself has evolved to nD NMR, with n typically <= 4.

The XML universe is representing another example where an instrument manufacturer or software provider is now given the choice of supporting XML, CML, ANIML or GAML or any combination thereof. The dilemma rests in the decision what to support. Acceptance in the community can often not be used as a criterion, as many of these formats and standards have little acceptance. The lack of acceptance in turn slows down development.

3:00 – Intermission

3:10 – 101 – Leveraging the AnIML specification for analytical data exchange. Stuart J Chalk, schalk@unf.edu.

Department of Chemistry, University of North Florida, Jacksonville, FL 32224, United States

The development to the eXensible Markup Language (XML) has opened up significant opportunities for storing scientific data. The Analytical Information Markup Language (AnIML) is an XML based specification for the storage of analytical instrument data that will make data exchange/interchange more uniform from both short-term and long-term (archival) perspectives. This presentation introduces AnIML and discusses the integration of AnIML to the research environment for storage and exchange of analytical data via i) the specification alone, ii) augmenting the specification with a research extension, and iii) embedding the specification in another XML format. Examples of how the research data can be integrated/searched will also be presented.


(1) Department of Chemistry, St. Olaf College, Northfield, MN 55057, United States, (2) Department of Chemistry, University of the West Indies, Mona, Jamaica

We present a simple extension to the JCAMP-DX format using two userdefined-data-labels, ##$MODELS and ##$PEAKS, to add 3D Jmol-readable models to the file and also associate spectral bands with specific IR and RAMAN vibrations, MS fragments, and NMR signals. The purpose of the JCAMP-MOL is to allow for a single file that can be read either by the standalone Jmol application (which now incorporates JSpecView) or by twin Jmol and JSpecView applets on a web page. Clicking on an atom or selecting an IR/RAMAN vibration in Jmol highlights a band or peak or fragment on the spectrum. Clicking on the spectrum highlights one or more atoms, starts an IR vibration, or displays an MS fragment in Jmol.

4:10 – Concluding Remarks

Wednesday, September 11, 2013

Before and After Lab: Instructing Students in ‘Non–Chemical’ Research Skills – PM Session
Indiana Convention Center
Teaching information skills to undergraduates is essential in the electronic age; it is very easy for anyone to publish information, making it extremely difficult to locate authoritative information. The University of Pennsylvania faculty wants all organic chemistry laboratory students to locate physical property values for substances used in lab, but, in two out of the three semesters that the course is offered, there is insufficient lecture time to devote to information skills training. The professor and the chemistry librarian addressed this issue using a series of small-group training sessions offered in the various lab sections during the first week of lab. The small class sizes allow time for discussion and better facility for hands-on experience of the resources discussed in class, and holding the sessions during the week of lab check-in allowed the librarian to devote a solid 1.5 hours to the training without taking time away from other class activities. This paper describes the outline of the class, desired learning outcomes, and reception of the class by teacher and students.

Unless a scientist is limited by their employer from exposing their scientific activities through publications and presentations, their future impact, whether expected to be at a bench, in front of an instrument or surrounded by robotics, will largely be represented online through their published works, their citation profile and other forms of recognition of their work by their peers. Search engines are already harvesting information about a scientist and aggregating into profiles such as those offered by Google Scholar Citations and Microsoft Academic Search. Rather than be limited to the online representation provided by such services students are encouraged to participate in the creation of their online profile and architect the representation of themselves online to as large a degree as possible to represent themselves to future employers and collaborators. This presentation will give an overview of potential approaches to participating in development of their online persona.

A series of high profile incidents related to health and safety in the academic chemistry laboratories have brought...
heightened public attention to this issue in the last few years. The United States Chemical Safety Board, the National Academies of Sciences, and the American Chemical Society, as well as legal authorities in specific jurisdictions, have all expressed concern about the current state of safety education in the laboratory sciences curriculum. This presentation will review these concerns; describe the information challenges associated with supporting laboratory workers, both in keeping themselves safe and documenting that they have followed Prudent Practices in pursuing their science; and provide an overview of the evolving role of environmental health and safety professionals in supporting a laboratory safety culture in academia.

2:35 – 106 – Other skills for post-graduates. Pamela J Scott, pamela.j.scott@pfizer.com

Department of Intellectual Property, Pfizer, Inc, Groton, CT 06340, United States

The skill section of the resume is gaining importance, but what goes there and how does one develop them? This talk discusses the skills that become part of one’s development outside of Chemistry and Chemical Information, and resources available to develop them. They include communication skills, written and oral, problem solving, time management, negotiation, setting priorities, working effectively on teams, the iterative process with clients, collaborating with peers, continuous learning, and fiscal responsibilities. The overlap of traditional learning and these soft skills has shifted the hiring criteria and value one brings as a potential employee.

Wednesday, September 11, 2013

Back to the Future: Print Resources in a Digital World – PM Session

Indiana Convention Center
Room: 141

Grace Baysinger, Organizers
Grace Baysinger, Presiding

3:10 pm – 5:30 pm

3:10 – Introductory Remarks

3:15 – 107 – Digitizing documents to provide a public spectroscopy database. Antony J Williams (1), williamsa@rsc.org, Colin Batchelor (2), William Brouwer (3), Valery Tkachenko (1).

(1) eScience and Cheminformatics, Royal Society of Chemistry, Wake Forest, NC 27587, United States, (2) eScience and Cheminformatics, Royal Society of Chemistry, Cambridge, United Kingdom, (3) Penn State University, Pennsylvania, United States

RSC hosts a number of platforms providing free access to chemistry related data. The content includes chemical compounds and associated experimental and predicted data, chemical reactions and, increasingly, spectral data. The ChemSpider database primarily contains electronic spectral data generated at the instrument, converted into standard formats such as JCAMP, then uploaded for the community to access. As a publisher RSC holds a rich source of spectral data within our scientific publications and associated electronic supplementary information. We have undertaken a project to Digitally Enable the RSC Archive (DERA) and as part of this project are converting figures of spectral data into standard spectral data formats for storage in our ChemSpider database. This
presentation will report on our progress in the project and some of the challenges we have faced to date.

3:45 – 108 – Whither the books? Managing access to print resources in an academic elibrary. Leah R McEwen, lrm1@cornell.edu.

Physical Sciences Library, Cornell University, Ithaca, New York 14853, United States

In the academic libraries of today and the future, the challenge is to increase access to literature and information while not losing quality resources previously collected. As libraries transition from print to electronic collections and branch service points consolidate, what happens to print materials and how do users (and librarians!) find what they need? Where are the books is a persistent question in the elibraries at Cornell, graduate students miss the classic texts, reserves are in high demand, and many specialized reference sets and journals have not yet appeared in the digital arena. We are tackling these challenges on multiple fronts, through enhanced discovery systems, consortial arrangements with peer institutions and a variety of order/print/digitize on demand type services. Do these strategies serve the needs of the users and the print collections? This presentation will consider input and feedback so far from a variety of projects.

4:15 – Intermission

4:30 – 109 – Keeping the books on campus: The University of Chicago approach to library collection. Andrea Twiss-Brooks, atbrooks@uchicago.edu.

John Crerar Library, University of Chicago, Chicago, IL 60637, United States

Like many of its peer institutions, the University of Chicago has grappled in recent years with the rapidly diminishing capacity for growing print collections. While the transition from print to electronic journals has slowed some aspects of that growth, the Library continues to collect significant amounts of printed materials. In addition, as a premier research library (the 9th largest in the U.S.) the Library is committed to retaining previously collected print materials, even when those materials may be available in online formats. To this end, the University undertook construction of the Joe and Rika Mansueto Library, a high-density shelving, automated storage and retrieval facility crowned by a stunning glass domed reading room and state of the art conservation and digitization laboratory and located at the heart of campus. The story of how the University decided to build this library in the age of digitization will be related.

5:00 – 110 – Challenges and opportunities for academic research chemistry collections in the 21st century. Grace Baysinger, graceb@stanford.edu.

Stanford University, United States

Building, managing, and preserving chemistry collections has become a careful balancing act. Due to competition for space, many science libraries have shrunk, merged, or closed. Because print resources are becoming “endangered species,” efforts by research libraries are underway to help insure that there is a critical mass in the number of copies preserved for future researchers. With the excitement and opportunities that digital collections offer, come new possibilities of enhancing discovery not only in online collections but also to print resources. This presentation will provide an overview of
the activities underway to provide current and long-term access to core chemistry resources at a major research university library.

Thursday, September 12, 2013

Exchangeable Molecular and Analytical Data Formats and their Importance in Facilitating Data Exchange – AM Session Data Standards to Support Publishing and Lab Notebooks

Indiana Convention Center
Room: 140

Cosponsored by COMP

Antony Williams, Robert Lancashire, Organizers
Robert Lancashire, Presiding

9:05 am – 10:40 am

9:05 – Introductory Remarks

9:10 – 111 – Importance of standards for data exchange and interchange on the Royal Society of Chemistry eScience platforms. Antony J. Williams (1), williamsa@rsc.org, Colin Batchelor (2), Jon Steele (2), Valery Tkachenko (1).

(1) eScience and Cheminformatics, Royal Society of Chemistry, Wake Forest, NC 27587, United States, (2) eScience and Cheminformatics, Royal Society of Chemistry, Cambridge, United Kingdom

The Royal Society of Chemistry provides access to a number of databases hosting chemicals data, reactions, spectroscopy data and prediction services. These databases and services can be accessed via web services utilizing queries using standard data formats such as InChI and moffiles. Data can then be downloaded in standard structure and spectral formats allowing for reuse and repurposing. The ChemSpider database integrates to a number of projects external to RSC including Open PHACTS that integrates chemical and biological data. This project utilizes semantic web data standards including RDF. This presentation will provide an overview of how structure and spectral data standards have been critical in allowing us to integrate many open source tools, ease of integration to a myriad of services and underpin many of our future developments.

9:40 – 112 – Extraction, analysis, atom-mapping, classification, and naming of reactions from pharmaceutical ELNs. Roger Sayle, roger@nextmovesoftware.com, Daniel Lowe, Noel O’Boyle.

NextMove Software, Cambridge, CAMBS CB4 0EY, United Kingdom

Electronic Laboratory Notebooks (ELNs) are widely used in the pharmaceutical industry for recording the details of chemical synthesis experiments. The primary use of this information is often for the capture of intellectual property for future patent filings, however this data can also be used in a number of additional applications, including synthetic accessibility calculations, reaction planning, and reaction yield prediction/optimization. Not only does a pharmaceutical ELN capture those classes of reactions suitable for small scale medicinal chemistry, but it is also uniquely a source of information on failed and poor yield reactions; an important class of data rarely found in the scientific literature or commercial reaction databases. This poster describes several of the technical chemoinformatics challenges in exploiting the wealth of synthetic chemistry
information in ELNs. Starting with the hand-drawn sketches stored in relational databases, we describe the steps required to transform and normalize this data into a clean and annotated reaction database in an “open” file format such as MDL’s RD and RXN formats, or reaction SMILES. This process includes the tricky steps of reaction atom mapping, role assignment of reactants, reagents, catalysts and solvents, and the recognition of a reaction as an example of a known named reaction (Suzuki coupling, Diels-Alder cyclization, nitro reduction, chiral separation etc.) Novel (and improved) algorithms for each of these tasks will be described, and where appropriate compared to and benchmarked against previous methods and implementations.

10:10 – 113 – How standards helped RSC to create The Merck Index Online. Richard Kidd, kiddr@rsc.org.

Royal Society of Chemistry, Cambridge, United Kingdom

We will talk about how The Merck Index* 15th Edition data was moved to RSC and converted using standards for text and structure data to create The Merck Index Online - making the data available online through a new platform within three months of receipt.


Thursday, September 12, 2013
Exchangeable Molecular and Analytical Data Formats and their Importance

in Facilitating Data Exchange – AM Session Data Exchange Standards in Chemistry and Drug Discovery

Indiana Convention Center
Room: 140

Cosponsored by COMP

Antony Williams, Robert Lancashire, Organizers
Antony Williams, Presiding

10:55 am – 12:00 pm

10:55 – 114 – Practical open data exchange formats for open organic chemistry projects. Jean-Claude Bradley (1), bradlejc@drexel.edu, Andrew SID Lang (2), Antony J Williams (3).

(1) Department of Chemistry, Drexel University, Philadelphia, PA 19104, United States, (2) Department of Mathematics, Oral Roberts University, Tulsa, OK 74171, United States, (3) Royal Society of Chemistry, Wake Forest, NC 27587, United States

This presentation will report on the author’s experience with Open Source tools and Open Data formats relevant to the execution of several organic chemistry and cheminformatics projects. These will include JCAMP-DX, SMILES, ChemSpiderIDs and related formats for representing spectra, collections of molecules, numerical data, reaction information and any other raw data necessary for the recording and analysis of organic chemistry experiments. It will be demonstrated that technical advancement in this area has far surpassed the state of cultural implementation in the academic community. It will be argued that a main point of inertia is likely the publishing
requirements for the majority of chemistry journals and that small changes could dramatically increase the amount of usable chemical information for the scientific community.

Bin Chen (1), Bing He (3), YING DING (2), dingying@indiana.edu, David Wild (3).

(1) Stanford University, United States, (2) Indiana University, United States, (3) Johns Hopkins University, United States

A critical barrier in current drug discovery is the inability to utilize public datasets in an integrated fashion to fully understand the actions of drugs and chemical compounds on biological systems. There is a need for both a resource to intelligently integrate the heterogeneous datasets pertaining to compounds, drugs, targets, genes, diseases, and drug side effects now available, and for robust, effective network data mining algorithms that can be applied to such integrative data sets to extract important biological relationships. Integrating heterogeneous data is clearly a prerequisite for important emerging research areas such as advanced network biology and network medicine. In this talk, we applied the Semantic Web (SW) technologies to integrate 25 public databases to facilitate drug discovery. We demonstrated the potentials of data mining and graph mining algorithms to identify hidden associations that could provide valuable directions for further exploration at the experimental level.

11:55 – Concluding Remarks

Thursday, September 12, 2013

General Papers – PM Session
Indiana Convention Center

Room: 140
Jeremy Garritano, Organizers
Jeremy Garritano, Presiding

1:00 pm – 2:30 pm

1:00 – 116 – Eureka Research Workbench: A semantic approach to an open source electronic laboratory notebook.
Stuart J Chalk, schalk@unf.edu.

Department of Chemistry, University of North Florida, Jacksonville, FL 32224, United States

Scientists are looking for ways to leverage web 2.0 technologies in the research laboratory and as a consequence a number of approaches to web-based electronic notebooks are being evaluated. In this presentation we discuss the Eureka Research Workbench, an electronic laboratory notebook built on semantic technology and XML. Using this approach the context of the information recorded in the laboratory can be captured and searched along with the data itself. A discussion of the current system will be presented along with the next planned development of the framework and long-term plans relative to linked open data.

1:30 – 117 – Chemotype approach to mapping the chemical landscape and exploring chemical-biological interactions within EPA’s ToxCast project. Rachelle J Bienstock (1), bienstock.rachelle@epa.gov, Chihae Yang (2), Jim Rathman (2), Ann M Richard (3).

(1) IS&GS - Civil, Contractor Supporting the EPA National Center for Computational Toxicology (NCCT) Office of Research &Development US Environmental Protection Agency, Lockheed Martin, Research Triangle Park, NC 27711, United States
U.S. EPA’s ToxCast project is employing high-throughput screening (HTS) technologies to profile thousands of chemicals that probe a wide diversity of biological targets, pathways and mechanisms related to toxicity. The current ToxCast chemical library is unprecedented in size (more than 1800 substances) and diversity, offering significant opportunities for cheminformatics contributions to toxicity modeling. However, the chemical diversity and nature of the HTS data sets present major challenges for QSAR modeling to contribute to the toxicity prediction problem. An approach employing a standard set of toxicity-informed feature sets, or chemotypes, is being developed to resolve the global chemical landscape into groupings of potential biological relevance. Use of these chemotypes in conjunction with biological knowledge and adverse-outcome pathway hypotheses, offers a means to focus and constrain modeling efforts into potentially productive areas of chemical and biological space, thereby improving modeling success and interpretability. Abstract does not represent EPA policy.
2013 CINF Officers and Functionaries

Chair
Dr. Antony Williams
VP Strategic Development
ChemSpider
Royal Society of Chemistry
919-201-1516 (voice)
williamsa@rsc.org

Chair Elect
Ms. Judith Currano
University of Pennsylvania
Chemistry Library
231 S. 34th St. 5th Floor
Philadelphia, PA 19104-6323
215-746-5886 (voice)
215-898-0741 (fax)
currano@pobox.upenn.edu

Past Chair/Nominating Chair
Dr. Rajarshi Guha
NIH Chemical Genomics Center
9800 Medical Center Drive
Rockville, MD 20852
814-404-5449 (voice)
812-856-3825 (fax)
rajarshi.guha@gmail.com

Secretary
Ms. Leah McEwen
Cornell University
Clark Library
283 Clark Hall
Ithaca, NY 14853-2501
607-796-6217 (voice)
607-255-5288 (fax)
leah.solla@cornell.edu

Treasurer
Dr. Rob McFarland
Washington University
Louderman Hall Room 549
1 Brookings Drive
Saint Louis, MO 63130-4862
314-935-4818 (voice)
314-935-4778 (fax)
mailto:rmcfarland@wustl.edu

Councilor
Ms. Bonnie Lawlor
National Federation of Advanced Information Services (NFAIS)
276 Upper Gulph Road
Radnor, PA 19087-2400
215-893-1561 (voice)
215-893-1564 (fax)
blawlor@nfais.org

Councilor
Ms. Andrea Twiss-Brooks
University of Chicago
4824 S Dorchester Avenue, Apt 2
Chicago, IL 60615-2034
773-702-8777 (voice)
773-702-3317 (fax)
atbrooks@uchicago.edu

Alternate Councilor
Mr. Charles Huber
University of California, Santa Barbara
Davidson Library
Santa Barbara, CA 93106
805-893-2762 (voice)
805-893-8620 (fax)
huber@library.ucsb.edu
Alternate Councilor
Dr. Guenter Grethe
352 Channing Way
Alameda, CA 94502-7409
510-865-5152 (voice and fax)
ggrethe@att.net

Program Committee Chair
Mr. Jeremy Garritano
Purdue University
M. G. Mellon Library of Chemistry
504 West State Street
West Lafayette, IN 47907-2058
765-496-7279 (voice)
jgarrita@purdue.edu

Membership Committee Chair
Dr. Gregory Banik
Bio-Rad Laboratories, Inc.
2 Penn Center Plaza, Suite 800
1500 John F Kennedy Blvd
Philadelphia, PA 19102-1721
267-322-6952 (voice)
267-322-6953 (fax)
gregory_banik@bio-rad.com

Archivist/Historian
Ms. Bonnie Lawlor
See Councilor

Audit Committee Chair
TBD

Awards Committee Chair
Ms. Andrea Twiss-Brooks
See Councilor

Careers Committee Chair
TBD

Chemical Information Bulletin Editor
Summer and Winter
Ms. Svetlana Korolev
University of Wisconsin, Milwaukee
2311 E. Hartford Avenue
Milwaukee, WI 53211
414-229-5045 (voice)
414-229-6791 (fax)
skorolev@uwm.edu

Chemical Information Bulletin Editor
Fall and Spring
Dr. Vincent F. Scalfani
The University of Alabama
Rodgers Library for Science and Engineering
109 Rodgers Library
Tuscaloosa, AL 35487-0266
205-348-5806 (voice)
205-348-2113 (fax)
vfscalfani@ua.edu

Communications and Publications Committee Chair
Dr. David Martinsen
American Chemical Society
1155 16th St NW
Washington, DC 20036-4801
202-452-2110 (voice)
d_martinsen@acs.org

Constitution, Bylaws & Procedures
Ms. Susanne Redalje
University of Washington
Chemistry Library
Box 351700
Seattle, WA 98195
206-543-2070 (voice)
curie@u.washington.edu
Education Committee Chair
Ms. Grace Baysinger
Stanford University
Swain Library of Chemistry & Chem. Engineering
364 Lomita Dr., Stanford, CA 94305-5006
650-725-1039 (voice)
650-725-2274 (fax)
graceb@stanford.edu

Finance Committee Chair
Dr. Rob McFarland
See Treasurer

Fundraising Committee Chair
Dr. Rajarshi Guha
See Past Chair/Nominating Chair

Tellers Committee Chair
Ms. Susan Cardinal
University of Rochester
Carlson Library
Rochester, NY 14627
585-275-9007 (voice)
585-273-4656 (fax)
escardinal@library.rochester.edu

Webmaster
Ms. Danielle Dennie
Concordia University
Vanier Library Building
7141 Sherbrooke St. W., Montréal (QC), H4B 1R6, Canada
514-848-2424 x 5237 (voice)
danielle.dennie@concordia.ca

Articles and Features
Mr. Bob Buntrock
Ms. Bonnie Lawlor
Ms. Song Yu
Dr. Svetla Baykoucheva

Committee Reports
Ms. Andrea Twiss-Brooks
Ms. Grace Baysinger

Sponsor Information
Mr. Graham Douglas
Dr. Rajarshi Guha

Technical Program
Dr. David Martinsen

Fall 2013 CINF Bulletin Contributors
## Schedule of Future ACS National Meetings

<table>
<thead>
<tr>
<th>Meeting</th>
<th>Date</th>
<th>Year</th>
<th>Location</th>
<th>Theme</th>
</tr>
</thead>
<tbody>
<tr>
<td>246th</td>
<td>Sept. 8–12</td>
<td>2013</td>
<td>Indianapolis, IN</td>
<td>Chemistry in Motion</td>
</tr>
<tr>
<td>247th</td>
<td>Mar. 16–20</td>
<td>2014</td>
<td>Dallas, TX</td>
<td>Chemistry and Materials for Energy</td>
</tr>
<tr>
<td>248th</td>
<td>Aug. 10–14</td>
<td>2014</td>
<td>San Francisco, CA</td>
<td>Chemistry and Global Stewardship</td>
</tr>
<tr>
<td>250th</td>
<td>Aug. 16–20</td>
<td>2015</td>
<td>Boston, MA</td>
<td>A History of Innovation: From Discovery to Application</td>
</tr>
<tr>
<td>251st</td>
<td>Mar. 13–17</td>
<td>2016</td>
<td>San Diego, CA</td>
<td>TBD</td>
</tr>
<tr>
<td>252nd</td>
<td>Aug. 21–25</td>
<td>2016</td>
<td>Philadelphia, PA</td>
<td>&quot;</td>
</tr>
<tr>
<td>253rd</td>
<td>Apr. 2–6</td>
<td>2017</td>
<td>San Francisco, CA</td>
<td>&quot;</td>
</tr>
<tr>
<td>254th</td>
<td>Aug. 20–24</td>
<td>2017</td>
<td>Washington, DC</td>
<td>&quot;</td>
</tr>
<tr>
<td>255th</td>
<td>Mar. 18–22</td>
<td>2018</td>
<td>New Orleans, LA</td>
<td>&quot;</td>
</tr>
<tr>
<td>256th</td>
<td>Aug. 19–23</td>
<td>2018</td>
<td>Boston, MA</td>
<td>&quot;</td>
</tr>
<tr>
<td>257th</td>
<td>Mar. 31–Apr. 4</td>
<td>2019</td>
<td>Orlando, FL</td>
<td>&quot;</td>
</tr>
<tr>
<td>258th</td>
<td>Aug. 25–29</td>
<td>2019</td>
<td>San Diego, CA</td>
<td>&quot;</td>
</tr>
</tbody>
</table>