Inside this issue:

Message from the Immediate Past Chair .................................................. 2
Letter from the Editor .............................................................................. 3
In Memoriam: Dr. William G. Town ......................................................... 4
Awards and Scholarships ........................................................................ 7
  2020 Herman Skolnik Award: Wendy Warr ......................................... 7
  Scholarship for Scientific Excellence ................................................... 8
  CSA Trust Grant .................................................................................. 9
Technical Program .................................................................................. 11
  Highlights of the ACS CINF Symposium on Extended Reality in Libraries .................................................. 11
Feature Articles ....................................................................................... 13
  Herman Skolnik Award Symposium 2019, Honoring Kimito Funatsu .................................................. 13
  New Directions for CINF: Strategic Plan, 2020-2023 ......................... 38
  CINF Luncheon: The End of a Tradition .............................................. 41
Book Reviews .......................................................................................... 44
Committee Reports .................................................................................. 45
  Council Report .................................................................................. 45
Sponsor Announcements ......................................................................... 50
  BioRad ............................................................................................... 51
  Thieme Chemistry ............................................................................. 52
CINF Officers .......................................................................................... 53
Contributors to this issue ......................................................................... 54
Dear Division of Chemical Information Colleagues,

It has been an honor to serve as chair of the Division of Chemical Information in 2019. Thanks to the work of CINF volunteers and committee chairs and the support of donors, the division had a productive year that included a number of initiatives, including the following.

**Membership survey**

With the support of ACS Research & Decision Support, we conducted a survey of CINF members. The goals of the survey were to determine what our members saw as strengths and priorities for our division and to gather their input on CINF’s performance. Overall, our members rated CINF as most effective at providing opportunities for exchange of technical information, such as meetings and symposia, and at providing networking opportunities; these were also functions that CINF members found highly important.

**Strategic planning**

We engaged in strategic planning in the fall of 2019, with the goal of re-energizing CINF. A strategic planning retreat, supported by a Committee on Divisional Activities Innovative Project Grant, provided the opportunity to create a roadmap to advance CINF’s mission. The three main thrusts of the 2020 strategic plan revolve around establishing a sustainable operational infrastructure for CINF, developing initiatives that engage and empower our members, and expanding our collaborations within and beyond ACS. Jeremy Garritano, 2020 CINF chair, is leading the implementation of the strategic plan and coordinating the work of several teams of volunteers.

**Program at National Meetings**

Technical programming is one of the core functions of our division, and it is highly valued by CINF members. In the spring and fall 2019 National Meetings we had a diverse and engaging program that included topics such as text mining and natural language processing, drug discovery, extended reality (XR) in libraries, Web-based chemistry databases, and machine learning in computational chemistry. The prestigious Skolnik Award symposium honored Prof. Kimito Funatsu for his contributions to structure elucidation, *de novo* structure generation and applications of cheminformatics methods to materials design and chemical process control.

Thank you for a fantastic year as chair of the Division of Chemical Information and for the opportunity to work with so many dedicated and talented CINF members. I look forward to continuing work with CINF under the various committees.

Elsa Alvaro  
CINF Immediate Past Chair  
elsa.alvaro@northwestern.edu
Greetings from unseasonably-warm Philadelphia!

I must begin this issue of the Chemical Information Bulletin (CIB) with an apology: I regret that this issue “hit the press” about two months later than I had planned and probably about three months later than it should have. Rather than offering excuses, I will simply state that it was a matter of balance. Balance is something that I have been actively seeking for many years, and I find that I am no closer to finding it than I was when I began my search. People frequently toss around the phrase “work/life balance”, usually referring to the way in which people grapple with the needs of their employers and their families, but one could just as easily struggle to balance two competing projects at work, or, in my case, fail to achieve a “teaching/running a library/editing a bulletin” balance, which puts me more in mind of a child’s mobile than a traditional pan balance. Let me apologize, therefore, for the fact that the editing segment of my mobile was off-kilter this winter and express my hope that you will find this issue of the CIB to be acceptable and maybe even worth the wait.

As is traditional, the winter issue of the CIB includes a detailed report of the symposium honoring the recipient of the Herman Skolnik Award. Once again, I would like to congratulate Prof. Kimito Funatsu on his award and on all of the exciting and innovative work that led to it. This year, however, we have an interesting plot twist: the article about the 2019 Skolnik Award Symposium was written by the 2020 Herman Skolnik Award recipient, Dr. Wendy Warr. Wendy has been a regular contributor to and proof-reader of the CIB for longer than I have been an editor, and, although we may not always agree on the optimal placement of commas, I love working with her. I am delighted to congratulate Wendy on this highly-deserved honor, resisting the urge to end my sentence with an exclamation point, which, I suspect, would not be in accordance with the ACS Guide to Scholarly Communication (formerly The ACS Style Guide).

Balance, however, demands that sorrow exist alongside joy, and, in this issue, we mourn the death of a long-time CINF colleague, leader, and mentor to many, Dr. Bill Town. I have known Bill for as long as I have been a member of CINF, and I suspect that I am not the first person to be surprised and flattered that such a distinguished scientist would take an interest in the career of an “imposter” like me. I will miss his smile, captured so beautifully in the picture printed alongside his obituary; his kindness; and the way in which he used his skills, influence, and connections to work to improve everything with which he came in contact.

2020 brings new beginnings, including an exciting new strategic plan for CINF, described by our incoming chair, Jeremy Garritano. The focus is truly on membership, with member benefits and communications forming a major component of the plan. 2020 also brings endings, though, including a fond farewell to the tradition of the CINF Luncheon. We reflect on past happenings as we revisit photos and technical symposia from the San Diego meeting, we examine the things that some of our sponsors are doing right now, and we look forward to the future. Perhaps you will want to curl up with a good book and wait for spring; our book reviewer in residence, Bob Buntrock, may have some suggestions for you!

I wish you all a happy, healthy, and, if possible, balanced 2020.

Judith N. Currano, Editor
University of Pennsylvania
currano@pobox.upenn.edu

Chemical Information Bulletin, 2019, 71 (4)
In Memoriam

Dr. William Geoffrey Town, March 31, 1943 to June 24, 2019

Our dear friend Bill Town passed away suddenly, of a cardiac arrest, on June 24, 2019, at the age of 76. Bill was the son of William Henry Town and Amy Town (née Morton). He grew up in Dagenham, Essex, U.K. He obtained a B.Sc. in chemistry from the University of Birmingham in 1964, and a Ph.D. from the University of Lancaster in 1967. His specialism was crystallography. Later, when he founded his own company, he achieved accreditation from the Chartered Association of Certified Accountants.

In 1967, Bill moved to Sheffield University, where he conducted research on chemical structure searching systems, as part of the pre-eminent research group led by Michael Lynch. In 1968, he moved to the University of Cambridge, where he worked under Olga Kennard in the Cambridge Crystallographic Data Centre, another world-leading group. He was an author on one of the first publications about the Cambridge Structural Database (Kennard, O.; Watson, D. G.; Town, W. G. Cambridge Crystallographic Data Centre. I. Bibliographic file. J. Chem. Doc. 1972, 12 (1), 14-19). CINF members will be interested to hear that Bill was attached to the school of librarianship.

In 1971, Bill moved to Ispra, in Italy, to work for the European Community Joint Research Centre. The family settled on the shores of Lake Maggiore, with views across the lake and the Alps in the distance (Bill loved sunshine), but they also had an apartment across the border in Switzerland. Bill became fluent in Italian: I loved to hear him speak this most musical of languages, but he also spoke French well, he spoke more German than I do, and later in life he also took up Spanish. Bill also got interested in skiing during his years in Italy.

At the Joint Research Centre, Bill led a team building the Environmental Chemical Data and Information Network (ECDIN). The team was also involved in the preparation of the European Inventory of Existing Commercial Chemical Substances (EINECS), which was to be published in seven languages. Bill was always an environmental crusader, but even more so in his seventies when he unfortunately developed chronic obstructive pulmonary disease. He was passionate about trying to reduce air pollution in London.

In 1983, the family returned to the United Kingdom, where Bill and his colleague, David Proctor, set up Hampden Data Services Ltd. (HDS). The HDS years were the height of Bill’s career. It was at HDS that STN Express was developed, and I had the pleasure of leading one of the key industrial teams (at ICI Pharmaceuticals) driving its development. The HDS chemical structure system, PSIDOM, was one of the first chemical structure editors, and an early way of building personal chemical structure databases and using structure entry as a front end to online searching systems.

In 1991, Bill relinquished his interest in Hampden Data Services (which still survives as a part of Chemical Abstracts Service), and founded William Town Associates, which he ran briefly, before joining Derwent Information (later Thomson Corporation) from 1992 to 1997, as Business Development Manager, Scientific Information. In 1997, he became Managing Director at ChemWeb, (later owned by Elsevier), and some readers may remember the Boston tea party reception that ChemWeb sponsored at a Boston ACS meeting. In 2002, Bill set up Kilmorie Consulting. This would later become Kilmorie Clarke, after he met Maggie Clarke and they decided to combine their businesses.
Bill was also a visiting professor in the Department of Information Studies at the University of Sheffield until 2006, and he was Chair of the Board of Governors at the Cambridge Crystallographic Data Centre, where his career had started. He was a member of the American Chemical Society for over 30 years, and he regularly attended ACS national meetings. He was a member of the CINF Publications Committee from 1993-1996, chaired the Awards Committee from 2002-2005, and chaired the Nominating Committee in 2001. He served the division as Chair-Elect in 1999, Chair in 2000, and Past-Chair in 2001, and was an alternate councilor from 2006-2008. He was the recipient of the 2008 CINF Meritorious Service Award.

Svetla Baykoucheva reports that Bill played an important role in the transition of the *Chemical Information Bulletin (CIB)* from print to digital. In 2009, he became chair of the CINF Publications Committee, and Svetla was editor of *CIB*. They organized the digitization of all print issues of the *Bulletin*, from its inception in 1949 to summer 2010. Bill secured a grant from ACS to pay for the scanning and the maintenance of the archive. He negotiated the price and conditions for access to the archive with the University of North Texas.

In the mid-1980s, Bill was secretary to what is now the Royal Society of Chemistry (RSC) Chemical Information and Computer Applications Group. He and Ian Tarr did all the organizational work involved in setting up the Chemical Structure Association Trust (CSA Trust), and Bill was a signatory to the Declaration of Trust on December 5, 1988. Bill was the first chairman (*sic*) of the trust.

Many people have spoken about the time and effort Bill put in helping to mentor them, providing them with his quiet influence, which helped them grow with the confidence to progress at the start of their careers. He was a great friend to many in this respect, and was always ready to help, and listen, particularly over a good meal and some fine wine and coffee. He was always full of new ideas and was constantly thinking of new things to do and pushing things in new directions.

In an interview that he gave to the CINF *Chemical Information Bulletin* (https://acscinf.org/content/career-chemistry-and-chemical-information) in 2009, Bill talked about his passion for exotic travels. He was always keen on travel, and, at the age of only 15, he cycled all the way from London to Cornwall. While at Cambridge, he attended a conference in Moscow and drove there with a friend in a hastily purchased and unreliable minivan, at a time when the cold war was in full swing and traveling behind the iron curtain was ill-advised. In 1999, he became an eclipse chaser after witnessing his first total solar eclipse in France. His love of the moment steered him to visit new and interesting places, including Botswana, Spain, Libya, Siberia, and Easter Island, to view more total eclipses. On a trip to Zimbabwe and Zambia, his canoe overturned and pitched him into the Zambezi on the first day, very shortly after he had sighted crocodiles. In 2007, he also toured Thailand and Cambodia, this time not chasing eclipses.

In later life, Bill became a political activist. He became a founding member of the Social Democratic Party (SDP) in 1981. Coincidentally, the announcement of the formation of the SDP was at St. Ermine’s Hotel in London, where large numbers of us were gathered for a major conference on chemical nomenclature. Bill followed the evolution of the SDP into the Liberal Democrats, as they are today. He was a passionate European and was committed to the European Union. He became a pivotal part of the Lewisham Liberal Democrats, and gave a lot of his time to serving on committees and helping to deliver leaflets, man stalls or help with information technology (IT) issues. In 2013, he even stood as a Liberal Democrat candidate for one ward. He was a keen marcher and took part in the many rallies and marches against Brexit. One of Bill’s happiest moments was the recent Lib Dem victories in the European elections, in areas which had never previously voted Lib Dem.

I knew Bill for more than forty years. We missed our first opportunity to meet in Sheffield in 1968, when he would have tutored me on a course on computer handling of chemical structural information. For some reason I did not attend; maybe the course was over-subscribed. So, we first met, briefly, at
a Chemical Notation Association conference in Kent in 1979. Our friendship blossomed after we both walked at the same time into the lobby of a hotel in Columbus, Ohio, in April 1982. Bill had come from Italy to visit Chemical Abstracts Service (CAS) on EINECS business, and, coincidentally, I was being given a tour of CAS on the same day. Our CAS hosts arranged for us to have lunch together.

Bill spoke about EINECS at the Fall 1982 ACS National Meeting in Kansas City, and I persuaded him to organize a symposium on EINECS at the fall meeting the next year, in Washington, DC. It was there that Bill showed me pictures of the house he was buying in Oxfordshire, England, and told me about the founding of HDS. That company played a big part of both our lives for the next eight years. During this period, Bill and I were also co-organizers of the first international chemical structures meeting, held in Noordwijkerhout in 1987. I reviewed Derwent CD ROMs when Bill worked for Derwent, and, from 1997 until 2002, I had a contract writing for ChemWeb.com. Space and time will not allow me to list all the projects we worked on together.

Bill was a very big part of my life for over 30 years. He was a source of business inspiration, a confidante who could be trusted with my secrets, an advisor when big decisions had to be made, and a shoulder to cry on when things were not going well. He was one of the cleverest people I knew. He was always able to stay calm, and do exactly the right thing in times of crisis. I loved to hear his voice.

As his son Matthew said, Bill was a gentleman, a scholar, modest, kind, sensitive, level-headed, quietly knowledgeable, always willing to help, a good man. Dear Bill, may you rest in peace. We offer our sincere sympathies to Maggie, Bill’s wife; to Anne, his first wife, and their children Helen, Matthew, and Amy; to his sister Pat Bollens; and to his grandchildren, Joe, Ella, Emi, Tom and Casper.

Wendy A. Warr
Wendy Warr & Associates
wendy@warr.com
Awards and Scholarships

2020 Herman Skolnik Award: Wendy Warr

Dr. Wendy Warr will be the recipient of the 2020 Herman Skolnik Award, presented by the ACS Division of Chemical Information, for her contributions to the fields of chemical information and a number of related fields that impinge on chemical information including chemical structure representation, substructure searching, retrosynthesis, and reaction prediction.

The prize consists of a $3,000 honorarium and a plaque. Dr. Warr will also be invited to organize an award symposium at the fall 2020 ACS national meeting, to be held in San Francisco, CA.

For more than 40 years, Dr. Warr has had a global influence on chemical information and cheminformatics. She provides services to the chemical information, cheminformatics, and computational chemistry communities worldwide, and has become a key opinion leader and trend watcher. In the early 1990s, she monitored such trends as combinatorial chemistry, chemistry and the Internet, and intranet and ethernet in industry. In the early 2000s, she switched focus to outsourcing and changing pharma R&D strategies, and she continues to keep track of today’s trends of AI, cloud computing, and blockchain.

Dr. Warr formed her current business, Wendy Warr & Associates, in 1992. Since that time she has successfully supplied business and competitive intelligence services to a broad spectrum of clients across the world. Her success stems from her extensive network, incredible energy, and deep curiosity, and her specialized market knowledge of chemical information, computational chemistry, drug discovery, cheminformatics, STM publishing, and scientific communication. Her clients have included at least 15 major pharmaceutical and chemical companies, venture capitalists, financial analysts, all of the well-known chemistry publishers, software companies, many cheminformatics and analytical chemistry companies, and many smaller commercial and not-for-profit companies and academic organizations. Scientific database producers have also benefited from her expert counsel and services in recent years.

Dr. Warr obtained her doctoral degree (D. Phil.) from the University of Oxford in 1971, and subsequently joined ICI Pharmaceuticals, where she held multiple positions, culminating in leading the department of Information Services. In 1992, Dr. Warr established her own company, Wendy Warr & Associates, which provides business and competitive intelligence services to a broad spectrum of clients in the United States, Europe, Australia, the Middle East, and Asia.

Dr. Warr has played key roles in several professional organizations including the American Chemical Society, the Royal Society of Chemistry, the German Chemical Society, the Society of Chemical Industry, the Chemical Structure Association Trust, and the Institute of Information Scientists. In many cases, she has been instrumental in shaping their activities. She has published over 80 articles in academic journals and over 100 commercial reports, along with numerous invited lectures at venues such as NIST (Washington DC) and the University of Strasbourg (France). She has been an associate editor for the Journal of Chemical Information and Modeling (as well as its predecessor, the Journal of Chemical Information and Computer Sciences). She has received numerous awards and honors, including the Ernie Hyde Award of the Chemical Structure Association (1984), and she is a Fellow of the Royal Society of Chemistry and a Fellow of the Chartered Institute of Library and Information Professionals.

Rajarshi Guha
Chair, CINF Awards Committee
awards@acscinf.org
Alzberta Tuerkova, winner of the CINF Scholarship for Scientific Excellence, receives the award from C4SE coordinator Stuart Chalk.

The 2019 Scholarship for Scientific Excellence was once again generously sponsored by ACS Publications. Here, Alzberta Tuerkova poses with Melissa Blaney of ACS Publications.

From left to right: Stuart Chalk (C4SE Coordinator), Alzberta Tuerkova, Melissa Blaney (ACS Publications), Elsa Alvaro (CINF Chair)
Applications Invited for CSA Trust Grants for 2020

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 now inviting the submission of grant applications for 2020.

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). Grantees have the option of payments being made in U.S. dollars or in British pounds equivalent to the U.S. dollar amount. 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. Proposals from those who have not received a grant in the past will be given preference. 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 (including the presentation of an already accepted research paper), to gain access to special computational facilities, or to acquire unique research techniques in support of one’s research. Grants will not be given for activities completed prior to the grant award date.

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 and a recent photograph.

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. A copy of the completed application document must be supplied for distribution to the Grants Committee and can be submitted via regular mail or e-mail to the committee chair (see contact information below).

Deadline for Applications
Application deadline for the 2020 grant is April 17, 2020. Successful applicants will be notified no later than May 25, 2020.
**Address for Submission of Applications:**
The application documentation can be mailed via post or emailed 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**

**2019**

*Vinicius Alves*, University of North Carolina Eshelman School of Pharmacy, Chapel Hill (U.S.A.), was awarded $2,572 to present his research paper entitled “Multi-Descriptor Read Across (MuDRA) as a novel computational approach for Chemical Toxicity Prediction” at the 10th International Symposium on Computational Methods in Toxicology and Pharmacology Integrating Internet Resources that was held in Ionnina, Greece, from June 23-27, 2019.

*Guilian Luchini*, Colorado State University, Fort Collins, CO (U.S.A.), was awarded $1,399.00 to attend the American Chemical Society meeting that was held from August 24-29 in San Diego, CA, where he presented his research in applying often-overlooked corrections to DFT frequency calculations in an automated fashion.

*Roi Rutenberg*, Chemistry Department, Stanford University, Stanford, CA (U.S.A.), was awarded $2,072 for travel to visit the University of Illinois, Chicago in order to model molecular dynamic (MD) simulations at the Kral group as part of his research related to retrieving information about pEtN cellulose’s chemical structure as an individual compound and as a partner in future chemical reactions.

*Monika Szabo*, Monash Institute of Pharmaceutical Sciences, Monash University, Victoria, Australia, was awarded $2,000.00 for travel to attend two conferences at which she presented her research on drugs for myelofibrosis. The conferences were EFMC-ASMC’19 International Symposium on Advances in Synthetic and Medicinal Chemistry, held in Athens, Greece, from September 1-5, 2019, and the 20th SCI/RSC Medicinal Chemistry Symposium, held in Cambridge, U.K., from September 8-11, 2019.

**2018**

*Stephen Capuzzi*, Division of Chemical Biology and Medicinal Chemistry at the University of North Carolina Eshelman School of Pharmacy, Chapel Hill (U.S.A.), was awarded a grant to attend the 31th ICAR in Porto, Portugal from 06/11/2018 to 06/15/2018, where he presented his research entitled “ComputerAided Discovery and Characterization of Novel Ebola Virus Inhibitors.”

*Christopher Cooper*, Cavendish Laboratory, University of Cambridge, U.K., was awarded a grant to present his current research on systematic, high-throughput screening of organic dyes for co-sensitized dye-sensitized solar cells. He presented his work at the Solar Energy Conversion Gordon Research Conference and Seminar held June 16-22, 2018 in Hong Kong.

*Mark Driver*, Chemistry Department, University of Cambridge, U.K., was awarded a grant to offset costs to attend the 7th EUCheMS conference, where he will present a poster on his research that focuses on the development and applications of a theoretical approach to model hydrogen bonding.

*Genqing Wang*, La Trobe Institute for Molecular Sciences, La Trobe University, Australia, was awarded a grant to present his work at the Fragment-Based Lead Discovery Conference (FBLD2018) in San Diego, U.S.A. in October 2018. The current focus of his work is the development of novel anti-virulence drugs, which potentially overcome the problems of antibiotic resistance of Gram-negative bacteria.

*Roshan Singh*, University of Oxford, U.K., was awarded a grant to conduct research within Dr. Marcus Lundberg’s Group at Uppsala University, Sweden, as part of a collaboration that he has set up between them and Professor Edward Solomon’s group at Stanford University, California. He conducts research within Professor John McGrady’s group at the University of Oxford. The collaboration will look to consolidate the experiments studies on heme Fe (IV)=O complexes currently being studied by Solomon’s group with future multi-reference calculations to be conducted within Lundberg’s group.
Technical Program

Highlights of the ACS CINF Symposium on Extended Reality in Libraries

On the morning of August 27, I was in for a treat. I attended “Extended Reality (XR) in Libraries & Beyond”, which I had co-organized with Michael Qiu and Nicolas Ruhs. The Carlson Science & Engineering Library, where I work, is planning to build a space to explore and teach about augmented reality (AR) and virtual reality (VR), so I was keen to learn about what was happening in other libraries.

Elisandro Cabado works in the Idea Lab at the Grainger Engineering Library at the University of Illinois (https://www.library.illinois.edu/enx/idealab/). He talked about why having XR in libraries makes sense. The area of augmented and virtual reality is rapidly evolving, and there is an exploding market. Libraries have always provided technology to explore information, and this is another way to share concepts. The library is heavily used by all types of students and is a place where people learn informally in public. Grainger already encourages creativity by providing crayons and cross-stitch. Why not offer 3D modelling? Emerging technologies are highly related to information literacy, and XR is being incorporated into research and instruction. Various disciplines are embracing it and making the learning experience more “viscerally engaging and immersive”. Libraries provide space, technology like Oculus and Vive, and expertise. For starters, Elisandro recommended play as a means of understanding a technology before attempting to integrate it into research or teaching. While students may be hesitant at first, once they put on the headsets, they do not want to remove them, and classes can run late.

When building an XR program, one should consider the following things:

- types of technology to offer
- setup time
- programming skills required
- where content can be obtained
- ways of overcoming users’ initial discomfort.


Tom Morrell is the Research Data Specialist at CalTech and works with Donna Wrublewski, Caltech’s Chemistry & Chemical Engineering Librarian. CalTech Library has a history of providing innovative services, including 3D Printing in the TechLab (https://www.library.caltech.edu/resources/techlab). They wanted to provide an affordable, small space entry point for XR, so, they positioned a VR workstation on a laptop cart. The user reserves an existing library study room and then folds the tables and chairs to clear the space. The key to the cart is then borrowed from the desk. Available content includes 3D printing files, and molecular models can be viewed and manipulated. The library offers a molecular modeling tutorial and uses Nanome (https://nanome.ai/), NASA’s Exoplanet Excursions VR experience (http://www.spitzer.caltech.edu/vr), and Google Earth (https://arvr.google.com/earth/), but it does not offer computer programming support. Projects are put in CalTech’s data repository. One example is VR Nucleosome Demo at https://doi.org/10.22002/d1.238. For more details, see https://www.library.caltech.edu/VR.

Sam Putnam (Engineering Librarian), Michelle Nolan (Chemical Sciences Librarian), and Ernie Williams from the Marston Science Libraries at the University of Florida described how they support XR. In MADE@UF (http://guides.uflib.ufl.edu/made), they provide equipment and training to students so that they can develop mobile VR and AR applications. They provide Unity (https://unity.com/), computers, headsets, Leaps, and similar technology, and they hold workshops and host speakers in their

Chemical Information Bulletin, 2019, 71 (4)
spaces. They also provide space for the GatorVR student club to meet and work with faculty to implement XR. One class focuses on using VR to solve social issues. The students first meet with people who are searching for ways to solve social problems and then build VR solutions. Two projects demonstrate ways to connect with chemical sciences students specifically. A 360° video was created to show proper safety techniques to new students. This is better than reading safety protocols because students learn about 3D aspects of safety procedures; for example, they can experience the exact physical methods of handling air-sensitive reagents, and it feels like they are actually in the lab. This preserves institutional memory and allows students to interact with the simulation at their own pace. Also, students can immerse themselves in a molecule to better understand it in 3D space. To see a video of this and to access the speakers’ full slide deck, go to https://bit.ly/UFatACS.

In the next talk, we experienced VR for a few minutes. A team of four, Rebecca Broyer (Associate Professor of Chemistry, University of Southern California), Steven Cutchin (Boise State), Sheree Fu (Engineering, Computer Science, and Technology Librarian, Cal State) and Dr. Shalini Ramachandran (Chemistry Librarian, University of Southern California), described how they created and implemented some training modules about laboratory safety. They played a demo of this experience so attendees could see how the students learned about spreading contamination from gloves. The student in the virtual lab was asked to perform some tasks while wearing gloves. Then, they were interrupted with printer problems and cell phone calls. Would they take the gloves off, or spread the contamination? After the scenario ended, the student could see all the places that they touched and where the chemicals were deposited.

In the final talk, Jodye Selco, a chemistry professor, talked about digital collections at the Cal Poly Pomona and California State University campuses. The emphasis there is on learning by doing, which is supported by an excellent collection called ScholarWorks that contains theses, dissertations, open access journals, and digitized scholarly materials. It is a shared institutional repository across 23 libraries.

In addition to the fine talks, the discussion was useful. The topics included getting started in Unity; ways to educate yourself, perhaps using YouTube and an inexpensive headset; ways that one can account for nausea among users of XR technology; and the continued abundance of digital memory. I was imbued with a new sense of confidence that our library was making a worthwhile transition. This orientation has helped me prepare for this change, and I was thrilled to meet and hear from these talented librarians and colleagues. Thank you for a pleasant and informative morning.

Susan Cardinal
University of Rochester
scardinal@library.rochester.edu

Are you interested in reliving your fall national meeting experience? Did you miss the national meeting but want to feel as though you were there?

Wendy Warr has posted photos from the fall ACS meeting in six albums at https://www.flickr.com/photos/cinf/albums. Enjoy!
Herman Skolnik Award Symposium 2019, Honoring Kimito Funatsu

A Report for the *Chemical Information Bulletin* by Wendy Warr

**Introduction**

Prof. Kimito Funatsu was selected to receive the 2019 Herman Skolnik Award for his contributions to structure elucidation, *de novo* structure generation, and applications of cheminformatics methods to materials design and chemical process control. His seminal contributions include the conceptualization and implementation of algorithms and expert systems for structure elucidation and chemical synthesis design, systems which have been extensively applied in the pharmaceutical industry. In recent years, he has increasingly focused on inverse QSAR analysis, including *de novo* structure generation, and the development of the soft sensor methodology for chemical process control. The latter approach represents another example of ground-breaking research with immediate practical and industrial application potential.

Kimito has secured large amounts of funding from the chemical and pharmaceutical industries to drive large-scale collaborative projects at the interface between academia and industry, most recently in the context of the CREST Program on Big Data Applications, funded by the Japan Science and Technology Agency. With more than 200 peer-reviewed publications, and a plethora of presentations and conference contributions, Kimito is among the core of leaders of the chemical information and informatics field worldwide.

He obtained his doctoral degree in physical organic chemistry from Kyushu University in 1983, and joined Prof. Shinichi Sasaki’s group at Toyohashi University of Technology in 1984. During his time with that group, he worked on a variety of cheminformatics applications including the structure elucidation system CHEMICS, the organic synthesis design systems artificial intelligence for planning and handling organic synthesis (AIPHOS) and knowledge base-oriented synthesis planning system (KOSP), and other systems in the areas of *de novo* design, and chemogenomics. In 2004, he moved to the University of Tokyo to continue research in these areas as a full professor, and there he expanded into material design and soft sensors for monitoring and controlling chemical plants. In addition to his professorship, he is the research director of the Data Science Center at the Nara Institute of Science and Technology (NAIST).

Kimito initiated the tradition of organizing biannual international cheminformatics schools in Japan. He also initiated the Computer-aided Chemistry Forum for scientific communication and practical training in cheminformatics, and established the Japanese Society of Cheminformatics. His relentless community service efforts also include his tenure as the President of the Division of Chemical Information and Computer Sciences of the Chemical Society of Japan (2004–2014). He has received several awards in recognition of his many contributions, including awards from the Japan Information Center of Science and Technology in 1988, from the Society of Computer Chemistry Japan in 2003, and from the Society of Chemical Engineering in 2017.

Kimito was invited to present an award symposium at the Fall 2019 ACS National Meeting in San Diego, CA. There were 10 speakers, in addition to Kimito himself.
Monitoring progress in lead optimization

Jürgen Bajorath of the University of Bonn presented a computational method termed Compound Optimization MOnitor (COMO)\textsuperscript{5} that helps to determine if further optimization progress can be expected for a given analogue series (AS) or if sufficient numbers of analogues have been generated. In COMO, virtual analogues (VAs) are used to populate the chemical space around an AS; chemical neighborhoods (NBHs) of analogues are defined, and VAs falling inside and outside NBHs are determined; potency distributions of analogues are analyzed; lead optimization relevant properties of analogues are evaluated; and multiple scores are calculated to quantify AS progression.

To produce VAs, the core of an AS is decorated with more than 16,000 substituents extracted from ChEMBL, applying 12 retrosynthetic (RECAP)\textsuperscript{6} rules. The VAs are sampled using RECAP-rule-compliant substituents and hydrogen atoms. A large number of analogue series from different sources have been studied, and alternative chemical space representations and virtual analogues of different designs have been explored.\textsuperscript{7}

Coverage of chemical space around ASs has been estimated by defining NBHs of experimental analogues and screening these NBHs with virtual compounds.\textsuperscript{7,8} To evaluate compound distributions in AS-centered chemical space and across NBHs of analogues, distances are calculated. The distance between two compounds is given by the Euclidean distance between two multidimensional vectors encoding molecular properties. VAs may or may not map to NBHs of existing analogues, and VAs may be located in overlapping NBHs (Figure 1).
To evaluate lead optimization progress, it must be determined how extensively chemical space around a given AS is covered, and how densely an AS samples the covered space (chemical saturation), and whether analogues display significant potency variations (SAR progression). Varying potency of structural analogues indicates SAR discontinuity.

A COMO scoring scheme was developed for profiling ASs that addresses the questions of chemical saturation and SAR progression. The chemical space coverage score ($C$) quantifies the VA coverage of all NBHs; the coverage density score ($D$) measures the VA coverage of overlapping NBHs; and the chemical saturation score ($S$) combines $C$ and $D$ ($2CD/C+D$).

The SAR progression score ($P$) quantifies potency variations of analogues sharing VAs in their NBHs. It is a VA-dependent measure of local SAR discontinuity. The SAR heterogeneity score ($H$) relates the potency distribution of analogues with overlapping NBHs to the mean potency of the AS. It is a VA-independent measure of global SAR heterogeneity. The multiproperty score ($M$) evaluates multiple compound properties. It is independent of the COMO scoring formalism, and includes a “traffic light” score for each individual ADME property of any active analogue.

Figure 2 shows exemplary analogues of a series of ATPase inhibitors, and virtual analogues falling into their NBHs. There are nine VAs, four NBHs, and three VAs in NBHs. Two of the VAs are in overlapping NBHs.

**Figure 1. Neighborhood analysis**

**Figure 2. Exemplary analogues belonging to a series of ATPase inhibitors (black), and their NBHs (light blue), and virtual analogues (red) falling into the NBHs.**

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coverage (3/9)</td>
<td>$C = 0.33$</td>
</tr>
<tr>
<td>Density 1-1/(6/3)</td>
<td>$D = 0.50$</td>
</tr>
<tr>
<td>Saturation</td>
<td>$S = 0.40$</td>
</tr>
<tr>
<td>SAR Progression</td>
<td>$P = 0.66$</td>
</tr>
<tr>
<td>SAR Heterogeneity</td>
<td>$H = 0.69$</td>
</tr>
</tbody>
</table>
Finally, Jürgen presented some unpublished work on analogue series profiling, in which his team studied 72 ASs, extracted from ChEMBL 24, with 1-6 substitution sites and 50-148 analogues. It was found that $S$ and $P$ scores are largely insensitive to varying VA counts. The standard VA population size was 3000 VAs. With increasing NBH radii, saturation increases and differentiates ASs (for 72 ASs and 3000 VAs). The standard NBH radius threshold (the distance threshold for the top 1% of closest pairwise distances between virtual analogues) was 1.0 (1st percentile). $P$ scores change very little because they largely depend on the potency distribution of existing analogues falling into overlapping neighborhoods. A plot of $P$ score against $S$ score, with points sized by AS size and colored by $M$ score, shows that $S$ and $P$ scores are uncorrelated and do not scale with AS size. It also shows that scoring differentiates between series with different chemical saturation and SAR progression characteristics.

COMO also includes a compound design component. VAs serve a dual role as diagnostic compounds and candidates for lead optimization. Virtual analogues generated for chemical saturation analysis provide a pool of candidates for synthesis. COMO can be combined with machine learning to produce predictive models for VA selection. Support vector regression is being used for potency prediction of VAs. The methodology is easily expandable to include multiple optimization of relevant properties. Practical applications are underway.

Electronic-structure informatics using 3D descriptors of molecules

Manabu Sugimoto of Kumamoto University began by paying tribute to Kimito Funatsu and other great scientists who have inspired us. He explained that electronic-structure informatics is a discipline which obtains chemical information from electronic structures and the responses of molecules. It has been applied by great scientists to structure-reactivity relationships, structure-activity relationships, and structure-property relationships.

Years ago, Manabu, working with Prof. Hiroshi Nakatsuji, and studied metal NMR chemical shifts with an ab initio molecular orbital method to establish a reliable method of calculating those shifts and to clarify electronic mechanisms and origins of the shifts. They found that both ground and excited states are important for quantitative structure-property relationships.

The Hubbard-Holstein model is a simple model to describe the behavior of solid crystalline materials by considering the action of electrons and phonons on a lattice. It is a combination of the Hubbard model (electrons on a lattice) and the Holstein model (phonons and their interactions with electrons on a lattice). The Hamiltonian for the Holstein model has the same first term as the Hubbard model (i.e., the electron hopping) but has two new terms concerning phonons. Electron-phonon coupling is important in the understanding of various chemistries, such as hole transport materials, organometallic chemistry, and organic light emitting diodes.

Both molecular structure and molecular properties are related to electronic structure, so, there is the possibility of doing cheminformatics as quantum chemists would. Energy is the secret. In a chemical reaction, the energy falls from the reactant state, then rises to the transition state, and then falls again as the products are formed. The theory of chemical equilibrium tells us that biological activity can be related to “energy changes”.

Until recently, most of the descriptors that Manabu and his colleagues have been applying correspond to spectroscopic features of molecules. This set of descriptors has limitations in describing three-dimensional features related to molecular recognition. Therefore, for efficient cheminformatics modeling, Manabu now suggests three dimensional descriptors, which represent topological features of interaction energy surfaces and molecular orbitals, and coarse-grained descriptions of three dimensional features of molecules (Figure 3).
Electronic states to be considered are neutral spin-singlet ground \((S_0)\) state; neutral spin-singlet excited \((S_n)\) states; lowest neutral spin-triplet excited \((T_1)\) state; ionized (cation) state; and electron-attached (anion) state. Molecular size is also important, for example, in ligand-target docking. The main contributors to intermolecular interaction energy are intermolecular distance, excitation energy, and transition dipole moment. Again, Manabu emphasized that both ground and excited states are important for quantitative structure-property relationships. Solvation-desolvation energy is important in ion exchange phenomena.

Manabu briefly described four applications of these 3D descriptors. Toshi Ideo has studied polyphenols as fatty acid synthase (FASN) inhibitors and has obtained a coefficient of determinants \((R^2)\) of 0.9098 between experimental and predicted \(\log IC_{50}\). For some terpene antibacterial reagents, an \(R^2\) of 0.763 was obtained for experimental versus predicted \(\log MIC\) (minimum inhibitory concentration). The descriptors were rather less successful in predicting the biological activity of some chemicals regulating food intake.

The fourth application, carried out by Alga Manggara, concerned acute aquatic toxicity of 33 alkylphenols. Toxicity is related to both chemical and physical descriptors. The model predicts the concentration of a substance that inhibits 50% of the growth \((IGC_{50})\) of a *Tetrahymena pyriformis* population within a designated period. A dataset from Cronin et al.\(^{13}\) was used. A correlation matrix for the 29 descriptors evaluated for the alkylphenols showed strong correlation between certain pairs. Some correlating descriptors were eliminated to give five descriptor sets with the following results:

- Electronic + Size \((R^2 = 0.950)\)
- Electronic + Size + (MO + Charge + Reactivity + Physical) version 1 \((R^2 = 0.929)\)
- Electronic + Size + (MO + Charge + Reactivity + Physical) version 2 \((R^2 = 0.946)\)
- Electronic \((R^2 = 0.798)\)
- Electronic + (MO + Charge + Reactivity + Physical) \((R^2 = 0.820)\)

Different regression models were obtained for the five sets because \(B, M,\) and \(EA\) (see Figure 3) are mutually correlated. The size of the alkyl group is important to the mechanism of action: smaller \(IGC_{50}\) leads to larger \(pIGC\), larger \(B\), larger \(M\) and smaller \(EA\) (lower LUMO).
In summary, Manabu described a new set of quantum chemical descriptors of molecules designed to describe three-dimensional features. He presented some applications showing reasonable correlations with experiments. In work on acute aquatic toxicity of alkyl phenols, the regression models were different for different descriptor sets because of strong correlation between descriptors.

**Fast evaluation of potential synthesis routes using DFT calculations on the basis of Transition State Database (TSDB)**

Four research groups are involved in a Japan Science and Technology Agency (JST) Core Research for Evolutionary Science and Technology (CREST) project “Development of a knowledge-generating platform driven by big data in drug discovery through production processes”. Makoto Taiji’s team at Riken first makes a very large scale virtual library (VLSVL) of drug candidate molecules, and adds synthetic routes and physicochemical properties. Yasushi Okuno’s group at Kyoto University takes information from the Riken group, studies ligand protein interactions, and passes back potential drug candidates to Riken. Kenji Hori of Yamaguchi University works on the rapid evaluation of the feasibility of synthetic routes, and shares routes and physicochemical properties with the Riken group. Kimito Funatsu’s group at Tokyo University is concerned with process control and the operation of chemical plants. They interact both with the Riken group and with Kenji’s group.

Kenji spoke about his own contributions. He described a procedure for *in silico* screening for synthesis routes. Systems such as Kimito’s transform-oriented synthesis planning system (TOSP), and knowledge base-oriented synthesis planning system (KOSP) usually offer many routes for any one target molecule. It may be hard to decide which route to try first and to confirm whether, in practice, the selected route actually produces the desired target. *In silico* screening addresses these problems.

Initially, organic chemists do a preliminary screening to select fewer than 10 potential reactions. These are the input for the *in silico* screening process, where transition state (TS) searches of the main and side reactions are carried out, and the Gibbs free energy of activation ($\Delta G^\ddagger$) is calculated. The reactions selected by this process are analyzed in more detail by estimation of solvent effects and study of the effectiveness of the route. The output is a set of ranked synthesis routes for experimental study. Kenji’s team has reported several successful implementations of this process.

The benefit of *in silico* screening is to exclude experiments which are unlikely to produce the target, dramatically decreasing the number of experiments needed and shortening the time spent on synthesis route development. Since it takes a long time to search TSs for reactions, Kenji aimed to shorten the CPU times for TS optimizations. Information on similar reactions is extremely useful in locating TSs, and, to obtain it, a database including TS information is needed.

Kenji’s team has gathered information on chemical reactions such as molecular names, keywords, optimized coordinates, and log files of quantum mechanical calculations and has constructed a Quantum Mechanical Calculations Results Database (QMRDB). The data in QMRDB are used to construct another database, the Transition State Database (TSDB). The PostgreSQL program is used for data handling, and the Open Babel program for molecular structure retrieval. The databases are searched in a Web browser.

The team has developed a cloud system for managing the databases and theoretical calculations. The user enters a SMILES string in a program called cStructure, on a Windows client, to search TSDB for mechanisms which use a similar reactant, or produce a similar product. The Tanimoto coefficient, TS coordinates, a chemical equation, and $\Delta G^\ddagger$ and $\Delta G$ values are returned by the database server in the cloud. iStructure, a program on the client, adds substituents at accurate positions and produces the
input for the Gaussian09 program, which is run on the server. A procedure has been developed in the iStructure program to create an initial structure for TS optimization for complicated targets, starting by downloading simple TS coordinates.

Kenji outlined how quantum chemistry assisted synthesis route development can contribute to synthetic chemistry in the 21st century. Synthesis planning systems usually offer many routes for a given molecule. In silico screening can drastically reduce the many possible routes to a few of the most feasible for experimental work. If the target is indeed obtained by actual synthesis, further work can be carried out, such as physicochemical property calculation. If the target is not obtained, feedback on the experimental results is passed to the in silico screening system.

In an example, target molecules from the VLSVL are sent to the Okuno team for deep learning predictions and docking calculations. The resulting drug candidates are passed to the in silico screening system. Ranked synthesis suggestions from that system are tested in synthetic experiments, and bioassays are carried out on compounds successfully synthesized. Active compounds are submitted for further investigation; data on inactive compounds are fed back into the docking system.

Molecules in the VLSVL were created by applying name reactions to molecules in a library of druglike molecules. Corresponding mechanisms were first examined to confirm whether or not TSs for the reactions existed. It is necessary to evaluate the toxicity of candidate molecules from the VLSVL using medicinal chemistry knowledge, and toxicity predictions are required to select potential candidates in the VLSVL and thus reduce the computational time needed for screening. It is very easy to construct initial structures of transition states using the iStructure program, but alternative synthesis routes have to be created when the synthesis route for the VLSVL is confirmed not to produce the target molecules. The TOSP and KOSP programs are used to suggest alternatives. Kenji presented two specific examples where the calculation of $\Delta G^\ddagger$ and $\Delta G$ for a transition state successfully supports the choice of alternative routes.

The design of new functional molecules is easy, but the development of their real synthesis routes is very difficult. Effort should be devoted to reducing the time wasted at this stage. Kenji gave three reasons why the synthetic routes from synthesis planning systems are not guaranteed to produce the targets: the precursors are much more complicated than the targets themselves; the number of synthesis routes diverges in multistep routes; and the reaction may not produce the target as the main product. A check based on cheminformatics can suggest a route that is likely to produce the desired target. It is possible to produce all the plausible products for a given set of reactants by using an appropriate reaction SMARTS.

Kenji closed by presenting an innovation cycle for developing functional molecules (Figure 4). AI molecular design suggests targets with the desired physicochemical properties and has a strategy for improving the functions of molecules. Synthesis routes from synthetic planning systems are produced but are not guaranteed to produce the targets. At the end of the cycle, results of the measurement of physicochemical properties of the targets are fed back into the AI system. The missing link is the connection of AI molecular design and in silico synthesis route development to the measurement of physicochemical properties of targets.
The missing link involves theoretical calculations, transition state searches, judgment on the basis of calculated $\Delta G^\ddagger$, and ranking synthesis routes. Experimental work can then confirm a synthesis route, and increase the yield of the reaction. The results of the experimental work are fed back into the theoretical calculation system. This whole subcycle, an innovative process for synthesizing targets, is the missing link which connects AI molecular design (followed by route development) to the measurement of physicochemical properties of targets.

Development using materials informatics in Japanese companies

The Materials Genome Initiative (https://www.mgi.gov) is a U.S. multi-agency initiative designed to create a new era of policy, resources, and infrastructure that support U.S. institutions in the effort to discover, manufacture, and deploy advanced materials twice as fast, at a fraction of the cost. When it was launched in 2011, the impact in Japan was significant, and similar efforts began in Japan. The discipline of materials informatics is perceived as new by some people, but Kimito Funatsu has long been working on the application of cheminformatics methods to materials design and chemical processing, and has laid the foundations of materials informatics.

Yukihiko Uchi of Asahi Kasei Corporation described work done in his own company in collaboration with Kimito’s team (http://www.mssj.jp/conf/62/program/2P-33.html). They have developed a structure prediction method for unknown compounds using two types of gas chromatography simultaneously plus mass spectrometry (GC x GC/MS) and quantitative structure-retention relationship (QSRR) inverse analysis models. Two-dimensional gas chromatography (GC x GC) is a gas chromatography technique that uses two different columns with two different stationary phases. The basic assumption of QSRR is that the retention time of GC has some correlation with various physical properties of compounds.\textsuperscript{18}

In forward analysis, a correlation model is built from known structures and their retention times. In inverse analysis, humans decide on the substructures that might represent key peaks in the MS of an unknown compound, and a structure generation program is used to generate plausible full structures. A correlation model is then used to predict the retention times for those structures, for comparison with the observed retention times. The quality of the decisions made about substructures depends on the experience and ability of the individual making the decision.
In work done by Yukihiko's colleagues, the first GC column has a nonpolar stationary phase (for boiling point separation), and the second column has a medium polarity phase (for polarity separation). There are two different retention times from the two columns. In forward analysis, molfiles of standard compounds are given Dragon6 descriptors (http://talete.mi.it/index.htm), the two retention times for each compound are measured, and a model is built for each type of retention time, using ensemble partial least squares regression. Excellent agreement was obtained between predicted and observed nonpolar retention times ($R^2 = 0.94$); the agreement was not quite so good for medium polarity retention time ($R^2 = 0.54$).

The structure generation algorithm used in inverse analysis is Chemish (http://www.cheminfonavi.co.jp/chemish/), developed in Kimito's laboratory. Molfiles for the candidate structures output by Chemish are given Dragon6 descriptors; a correlation model is used to predict the retention times for those structures; those retention times are compared with the measured ones; and candidate structures with times close to those of the unknown compound are ranked in order of probability.

Yukihiko presented two verification examples. The first is shown in Figure 5. The correct candidate was ranked 23rd among 359 structures and had retention times of 36.7 minutes and 5.2 seconds. (The first and second choices had retention times of 37.6 minutes and 4.7 seconds; and 37.1 minutes and 5.0 seconds.)

A second verification example is shown in Figure 6. There were 72 candidate structures. The correct candidate was ranked third with retention times of 47.0 minutes and 4.8 seconds. (The candidates ranked first and second had retention times of 44.7 minutes and 4.5 seconds; and 46.3 minutes and 4.5 seconds.)
In this study, QSRR is a technology that has potential but still needs improvement. The accuracy of the prediction model is not good enough when using the medium polarity column, and the accuracy of the substructure generation depends on the ability of an individual.

In materials development at Asahi Kasei Corporation, 10 types of raw materials have been selected from 80 possible types and, until now, have been advanced based on human intuition and experience. There are innumerable combinations to determine the ratio of the amounts of materials. It is possible to reduce the number of trials and errors by using informatics technology such as the inverse QSAR technology described here. Using cheminformatics is a response to the era of data-driven material informatics.

**Prediction and control of a vacuum deposition process by a data-driven method**

Yoichi Zushi and Yuya Takeda of Kaneka Corporation discussed two different examples: the development of a soft sensor in a thin-film photovoltaic (PV) deposition process, and the development of a prediction and control method of an organic, light-emitting diode (OLED) deposition process.

A thin-film solar cell (Figure 7) is made by depositing one or more thin layers of photovoltaic material on a substrate, such as glass, plastic, or metal. Amorphous silicon is a non-crystalline, allotropic form of silicon and the most well-developed thin film technology to date. A new attempt to fuse the advantages of bulk silicon with those of thin-film devices is thin-film polycrystalline (PC) silicon on glass. These types of thin-film cell are mostly fabricated by a technique called plasma-enhanced chemical vapor deposition (CVD). Post-processing such as laser treatment or sputtering follows.
PV efficiency greatly depends on the deposition conditions of silicon layers. Material gas flow rate, chamber pressure, and radio frequency power are each controlled by a device, but substrate temperature is indirectly adjusted with a heater and is affected by CVD operating status. The temperature falls sharply as polycrystalline silicon follows amorphous silicon. It is important to control the substrate temperature during deposition. In practice, the substrate temperature is difficult to measure online because there is a vacuum inside the process chamber, and there is no space for sensor installation, etc. A soft sensor was thus developed.

The modeling conditions are as follows. The response variable is substrate temperature. The data are acquired by pasting a thermocouple on the substrate, connecting it with the temperature logger, inputting it into the process chamber, and following the temperature during a few hours for the deposition batch. Explanatory variables (about 300 of them) are related to the dynamic characteristics of measurable material gas flow rate, chamber pressure, radiofrequency power, and panel heater temperature. Each process data item affects the substrate temperature with a time delay. Substrate temperature data are acquired under changing deposition conditions.

The first prediction model was created based on the time from the start of the batch, but an accurate prediction model could not be obtained due to the small amount of data. The data were therefore divided according to the features of each of the process steps, and multiple prediction models were created. The modeling method was partial least squares. The substrate temperature was then predicted accurately over time. Excellent agreement was obtained between predicted and observed values ($R^2 = 0.993$, RMSE = 0.90).

An online monitoring system was developed for the predicted substrate temperature. In short, the team developed a substrate temperature soft sensor during deposition with a small amount of data. The substrate temperature is predicted online, and the substrate temperature is stabilized by monitoring and control.

The second part of this presentation concerned a prediction and control method for an OLED device deposition process. The performance of OLED devices depends on the deposition process of organic materials. More than 24 hours are needed from the start of the deposition process until quality inspection, so quick feedback is prevented. There are quality specifications such as brightness, driving voltage, and color etc., but they have a trade-off relationship, so it is difficult to decide on operating conditions. The researchers developed a prediction and inverse analysis method to decide on operating conditions. The objective was to take the data from the quality inspection process, and feed them into an AI-based system to produce quality prediction and monitoring data and operating conditions that could be used in the next organic materials deposition experiment; the cycle then could be reiterated.

In the model, the explanatory variables ($x$), such as temperature and pressure, are a function of the response variables ($y$), such as voltage and brightness. Methods such as partial least squares, gradient boosting, support vector regression, ElasticNet, and random forest were used to build the model. In the inverse analysis, conditions ($x$) that satisfy the quality requirements ($y$) with the model were obtained, as were constraints on the conditions such as boundary and time. Multiobjective optimization using a genetic algorithm was then used to propose new operating conditions.

Machine learning models were obtained that were sufficiently accurate. The speaker showed good straight line plots for predicted versus experimental values for a number of $y$ variables in training and test set prediction. In the inverse analysis and multiobjective optimization using a genetic algorithm, if the manipulated variables were randomized all at once, in some cases, the researchers failed to obtain values of $x$ when $y$ satisfied the requirements. This was because the OLED lighting device has a layered structure (anode, organic layer, lighting layer, organic layer, lighting layer, organic layer, cathode), and there are strongly related variables in each layer. A new method was therefore used where...
the variables were grouped by layer. This method worked much better. The core technology of this system can be used in other cases, and the prediction and inverse analysis method will be used for other processes in future work at Kaneka.

**Designing synthesizable, bioactive compounds with chemistry-savvy machine intelligence**

Gisbert Schneider of ETH Zurich, Switzerland, recently co-authored a paper with Kimito Funatsu and others, which stated that QSPR and QSAR are shifting from a mere prediction of property or activity towards design. Gisbert started his talk by summarizing three approaches to the issue in drug design of “what to make next”. The chemist uses expert knowledge and intuition, with the knowledge represented both explicitly and implicitly. The “rational machine” uses rules and chemical transformations, where the knowledge representation is explicit. The “intuitive machine” uses distribution sampling, with implicit (probabilistic) knowledge representation.

Gisbert’s first *de novo* design approach for small molecules, the TOPology-Assigning System (TOPAS), was based on (al)chemical transformations. The World Drug Index was fragmented by RECAP into 25,563 unique building blocks which could be recombined to make new molecules.

Later Schneider’s team developed a reaction-based *de novo* design system, Design of Genuine Structures (DOGS). The compound construction procedure explicitly considers compound synthesizability, based on a compilation of 25,144 readily available synthetic building blocks and 58 established reaction principles. This enables the software to suggest a synthesis route for each designed compound. A combinatorial explosion in the structure generator is prevented by machine learning models, heuristics, and intuition. DOGS has been used successfully in the *de novo* design of small molecules as natural product mimetics. Further applications of DOGS have been published recently.

In an article in the Toronto National Post published on May 30, 2019, and updated on June 6, 2019, Joseph Brean quotes David Gunkel, a philosopher of robotics and ethics at Northern Illinois University. Gunkel said “We are now at a point where we have AI [systems] that are not directly programmed. They develop their own decision patterns.”

Very recently Gisbert’s team has reported a method for *de novo* design that uses generative recurrent neural networks (RNN) containing long short-term memory (LSTM) cells. This computational model captured the syntax of molecular representation in terms of SMILES strings with close to perfect accuracy. The SMILES strings were from compounds in ChEMBL with nanomolar activity. The “deep-learned” pattern probabilities can be used for *de novo* SMILES generation by fragment growing. This molecular design concept eliminates the need for virtual compound library enumeration. By employing transfer learning, the general RNN model was fine-tuned on recognizing retinoid X and peroxisome proliferator-activated receptor (PPAR) agonists. Five top-ranking compounds designed by the generative model were synthesized. Four of the compounds revealed nanomolar to low-micromolar receptor modulatory activity in cell-based assays. Apparently, the computational model intrinsically captured relevant chemical and biological knowledge without the need for explicit rules.

A very recent development is a bidirectional RNN-LSTM model. In the past, SMILES strings were generated from the generation point towards the right; now the method works as in Figure 8. The novelty of valid SMILES generated was 92 ± 2 % for the unidirectional RNN-LSTM model, and 97 ± 3 % for the bidirectional RNN-LSTM model.
Schneider’s team is now exploring the capability of “chemistry-savvy” machine intelligence to generate synthetically accessible molecules in Design of Innovative NCEs Generated by Optimization Strategies (DINGOS). This is a virtual assembly method that combines a rule-based approach (predicting building blocks for synthesis) with a machine learning (neural network) model trained on successful synthetic routes described in chemical patent literature. This unique combination enables a balance between ligand similarity based generation of innovative compounds by scaffold hopping and the forward-synthetic feasibility of the designs. In a prospective proof-of-concept application, DINGOS successfully produced sets of de novo designs for four approved drugs that were in agreement with the desired structural and physicochemical properties. Target prediction indicated more than 50% of the designs to be biologically active. Four selected, computer-generated compounds were successfully synthesized in accordance with the synthetic route proposed by DINGOS. The results of this study demonstrate the capability of machine learning models to capture implicit chemical knowledge from chemical reaction data, and suggest feasible syntheses of new chemical matter.

Valid chemical structures with explicit synthesis routes are produced, in a synthesizable chemical space, with implicit reactivity scoring. The reaction forecasting involves node expansion with DINGOS, navigation with Monte Carlo Tree Search, and use of a reward, or “scoring function” (e.g., similarity to the template, or activity prediction). Gisbert coined the term DinGO in order to allude to the game “GO” played in DeepMind’s Alpha Go Zero. DinGO plays the “what if?” game.

Another of Gisbert’s projects uses deep convolutional neural networks (CNNs). His team has published a hybrid CNN approach for molecular pattern recognition in drug discovery. Using self-organizing map images of molecular pharmacophores as input, CNN models were trained to identify C-X-C chemokine receptor type 4 (CXCR4) modulators with high accuracy. The machine learning classifier identified first-in-class, synthetic CXCR4 full-agonists. Additional macromolecular targets of the small molecules were predicted in silico and tested in vitro, revealing modulatory effects on dopamine receptors, and chemokine receptor type 1 (CCR1). These results positively advocate the applicability of molecular image recognition by CNNs to ligand-based, virtual compound screening, and demonstrate the complementarity of machine intelligence and human expert knowledge.

Gisbert concluded with some comments on the applicability of AI. We can expect several things from AI-driven molecular design: readily synthesizable, inspiring designs; similar success with rule-driven and data-driven AI; and better decisions for “failing early” and “choosing wisely”. We cannot expect drugs from scratch, or flawless prediction models.
Kiyoshi Hasegawa of Chugai Pharmaceutical Co. and his colleagues have applied an activity landscape technique to mouse, rat and human clearance data in order to select lead compounds from huge, high-throughput screening datasets. A naïve Bayesian method with ECFP_6 fingerprints from Pipeline Pilot was used. The frequencies of active and inactive states for each substructure were counted, and structural descriptors with biased frequency were selected. An in-house set of mouse clearance data was used. Mouse liver microsome clearance, $\text{CL}_{\text{int}}$, was measured (in $\mu$L/min/mg protein). The dataset was split into 25,000 assay results for the training set and 1000 for the test set. The threshold between stable and unstable was set to 30 $\mu$L/min/mg protein to maximize power of discrimination between stable and unstable. The prediction accuracy for the test set was 78% (true 76%, false 80%).

Activity landscape representations of different types of compound sets were calculated from potency data and pairwise compound distances in chemical space. From the fingerprints, a coordinate-free chemical reference space was generated by calculation of pairwise compound distances (dissimilarities). The set of all pairwise distances defines this reference space. Then, multidimensional scaling was used to project these molecules from the coordinate-free reference space onto an $x/y$-plane on the basis of the chemical dissimilarities. The 2D map was then color-graded by a geographic method.

The prediction model was applied to compounds that had already been through high throughput screening, including 22 classes and 819 actives. The mouse stability of all of the primary actives was predicted, and chemical classes were identified in which all members of the class were predicted to be unstable. In this experimental validation, the result was 97% correctness for unstable compounds. These compounds fell into seven classes with no stable compounds.

Kiyoshi has built a website where the structures of up to three species can be input for prediction of stability or instability, with a prediction score. A structure is displayed, colored to show the metabolically labile and stable atoms. A talk about this was given by J.T. Metz et al. at the SciTegic Users Group Meeting in 2007. The colored activity landscape can be viewed with Pipeline Pilot, using an interactive link from a circle in a scatter plot to see the chemical structure and its data.

Another issue is the gap between the enzyme and cell activities. This phenomenon is often encountered in drug discovery: the cell activity of the molecule is not high even though the enzyme activity is high. Comparing two activity landscapes of the enzyme and cell activities, it is possible to investigate which molecular skeleton is a promising target for lead optimization. Since the promising chemical space can be easily detected, libraries to fill that space can be designed. Kiyoshi has constructed a prediction model for cell activity from enzyme activity, and log$D$ models using Simulations Plus AD-MET Predictor (https://www.simulations-plus.com/software/admetpredictor/). He showed activity landscape displays for two selected molecules (Figure 9).
Drug repositioning involves identification of new therapeutic effects of existing drugs and of compounds that failed to be approved in the past. It is an efficient strategy for drug development that has attracted much attention. A great deal of information on existing drugs is available (e.g., information on safety, manufacturing processes, and pharmacokinetics). Drug repositioning can increase the success rate of drug development, and reduce the cost in terms of time, risk, and expenditure. A well-known example is sildenafil (Viagra), which was developed as a treatment for angina but was repositioned to treat erectile dysfunction and pulmonary hypertension.

Yoshihiro Yamanishi at the Kyushu Institute of Technology, and his colleagues, have developed novel machine learning methods that can be used to predict new associations between drugs and diseases, based on the molecular understanding of a variety of diseases: disease-causing genes, disordered pathways, environmental factors, and abnormal gene expression. Characteristic molecular features are often shared among different diseases. For example, the abnormal expression of phosphodiesterase type 5 (PDE5) is observed in both erectile dysfunction and pulmonary hypertension. Networks of drug-disease relationships can be produced by machine learning methods based on molecular features of drugs and diseases.

Yoshihiro has proposed a pathway-based drug discovery approach. A traditional approach is to search for drugs that regulate a single biomolecule, but, in this approach, molecular interactions between biomolecules are not taken into account. In pathway-based drug discovery, the approach is to search for drugs that regulate a pathway; molecular interactions are considered by using pathway information. Integration of drug-induced gene expression data with molecular network analysis can lead to prediction of new therapeutic effects of drug candidates.

Activated and inactivated pathways are identified from drug-induced gene expression signatures. The up- and down-regulated genes in the signatures are mapped onto many biological pathway maps, and the enrichment of the up- and down-regulated genes in each pathway is evaluated by pathway enrichment analysis. The 163 biological pathways in the Kyoto Encyclopedia of Genes and Genomes (KEGG, https://www.genome.jp/kegg/) were used. Now, $z = |G_{drug} \cap G_{pathway}|$ where $G_{drug}$ denotes a
set of up- or down-regulated genes in a signature induced by a drug, and \( G_{\text{pathway}} \) denotes a set of genes in a pathway map. Assuming that \( z \) follows a hypergeometric distribution,\(^{36}\) the probability of observing an intersection of size \( z \) between \( G_{\text{pathway}} \) and \( G_{\text{drug}} \) is computed as in Figure 10. The gene expression values in the signature of each drug are represented with a feature vector.

\[
P(\text{P-value}) = \sum_{i=0}^{\min(k, r)} \binom{k}{i} \binom{l-k}{r-i} \binom{l}{i}
\]

**Figure 10. Pathway enrichment analysis**

Yoshihiro has collaborated\(^{36}\) with Kenzaburo Tani, at the University of Tokyo, on pathway-based drug discovery for cancers. They analyzed chemically induced gene expression data of 1112 drugs on 66 human cell lines and explored drugs that inactivate cell cycle pathways, activate p53 signaling pathways, and activate apoptosis pathways. They performed a large-scale prediction of potential anticancer effects for all the drugs and experimentally validated the results. They successfully identified several potential anticancer drugs.

Natural medicine (e.g., Kampo in Japan) is popular, but the mechanism of action of these treatments is unclear. In “ordinary” medicine the mode-of-action is based on the interaction of one compound with one target. In natural medicine, multiple compounds interact with multiple targets, and the target proteins may work cooperatively. Perhaps pathway analysis on compound-induced transcriptome data would be helpful in such cases.

There have been numerous publications\(^{37-43}\) recently on identification of the modes of action of drugs, and prediction of drug therapeutic indications. It is very difficult and expensive to observe gene expression profiles experimentally for all combinations of drugs and human cell lines, so large parts of drug-induced gene expression data are unknown or unobserved. Connectivity Map (CMap), in which genes, drugs, and disease states are connected by virtue of common gene-expression signatures, was scaled up, as part of the NIH Library of Integrated Network-Based Cellular Signatures (LINCS) Consortium.\(^{44}\)

A novel gene expression profiling method, L1000,\(^{44}\) was used in the LINCS program, and it has opened the door to the large-scale analysis of drug-induced transcriptome data (drug profiles). However, there are far more unknown or unobserved values than known ones. This can be connected to disease profiles: disease-specific transcriptome data on highly and lowly expressed gene profiles. Alzheimer’s disease, asthma, atopic dermatitis, breast cancer, cystic fibrosis, inflammatory bowel disease, dengue, adrenoleukodystrophy, and many more diseases have been studied. For example, Wang et al. have reported a crowdsourcing project to annotate and reanalyze a large number of gene expression profiles from Gene Expression Omnibus (GEO).\(^{45}\) A cleaned database of extracted signatures was used to visualize and analyze these signatures on the CRoad Extracted Expression of Differential Signatures (CREEDS).

Previous methods for missing value imputation or data completion\(^{46-52}\) are applicable to matrix-structured data. Yoshihiro and co-workers have proposed a method applicable to tensor-structured data: Tensor-Train Weighted OPTimization (TT-WOPT).\(^{53}\) They applied TT-WOPT to drug-induced transcriptome data: 16 cell lines, 261 drugs, and 978 genes represented by a 261 \( \times \) 978 \( \times \) 16 tensor. As a baseline method, they also tested the CP-WOPT algorithm,\(^{54}\) which is a previously established tensor decomposition method applicable to data completion tasks. In the cross-validation experiments for performance evaluation, they randomly added artificial missing values to the original data before imputation. The relative standard error (RSE) between the original tensor and the one with imputed values was measured. In the case of artificial missing rates of 10% for the cell lines in total, RSE for
TT-WOPT was 0.0694 compared with RSE = 0.0750 for a nearest neighbor approach. In only one of the cell lines was the RSE for the nearest neighbor method (0.0415) better than RSE for TT-WOPT (0.0416). TT-WOPT also works well for 50% and 90% missing rates.

In Yoshihiro’s work, the original drug-induced transcriptome data are subjected to tensor decomposition to get a new version, including imputed data. From the latter a drug indication prediction can be made. For comparison, three existing transcriptome-based drug repositioning methods were compared, with and without tensor decomposition: inverse signature, XSum, and multitask learning. A benchmark dataset of 353 associations (261 drugs and 46 diseases) was used. The area under the receiver operating characteristic curve (AUC) was measured. Tensor decomposition contributed to more accurate prediction of drug indications in most cases. AUCs were more than twice as large for the multitask learning method for all 16 cell lines. Moreover, tensor decomposition is more effective in cell lines with high missing rates.

Yoshihiro gave two examples of predicted indications which have been confirmed with independent resources. Amodiaquine is an antimalarial drug. A predicted indication was pituitary adenomas, and this was confirmed by the literature. Niclosamide was originally an anthelminthic. A predicted indication of adult T-cell leukemia was confirmed by the literature.

Yoshihiro concluded that machine learning methods can predict new therapeutic effects of drug candidate compounds. Pathway analysis is useful for mode-of-action identification and drug discovery, and tensor decomposition for omics data contributes to enhancing the performance of drug indication prediction. Such methods will speed up the delivery of necessary drugs to patients.

Integrated cheminformatics and bioinformatics data science

Shigehiko Kanaya of the Nara Institute of Science and Technology and his colleagues have constructed a species-metabolite database for plants, the KNApSAcK Core Database, which contains (as of April 2019) 51,179 metabolite entries, 22,944 species entries, and 116,315 metabolite-species pair entries. This sort of database is useful because it allows the systematic analysis of large numbers of organic compounds with known and unknown structures in metabolomics. Shigehiko’s team has also developed a search engine for the database, making it possible to search for metabolites based on an accurate mass, molecular formula, metabolite name, or mass spectrum in several ionization modes. Various other databases can also be accessed on the KNApSAcK website (http://kanaya.naist.jp/KNApSAcK_Family), and the search engine can be downloaded (Figure 11).
Shigehiko believes that data science can be created by integrating disciplines including theoretical understanding in various research fields, informatics (for data systemization, model construction, and prediction), and statistics (for validation). Specifically, chemistry and chemical physics could be overlapped with chemical information (molecular structures), and chemometrics validation. An example is the overlap of deuterium isotope effects in solvolysis reactions, AlPHOS, and computer-aided structure elucidation (CHEMICS and soft sensors). Kimito Funatsu sits at the center of the overlap of such systems.

The role of data science is in moving from vertical relationships to horizontal relationships (Figure 12), standardizing mining techniques. The KNApSAcK family (Figure 11) is an example, allowing the understanding of biology based on natural products databases.

There are about 20,000 alkaloids in the KNApSAcK database, but few of their biosynthesis pathways are fully identified. Shigehiko and his co-workers have constructed a model to predict the precursors of alkaloids based on multi-graph convolutional neural networks (MGCNN). It is sometimes difficult for current fingerprint representations to emphasize specific features for target problems efficiently. It is advantageous to allow the model to select the appropriate features according to data-driven decisions. By encoding a molecule as an abstract graph, applying "convolution" on the graph, and training the weight of the neural network framework, the neural network can optimize feature selection for the training problem. By incorporating the effects from adjacent atoms recursively, graph convolutional neural networks can extract the features of latent atoms that represent chemical features of a molecule efficiently. The researchers trained the network to distinguish the precursors of 566 alkaloids, which are almost all of the alkaloids with known biosynthesis pathways, and showed that the model could predict starting substances with an average accuracy of 97.5%. The prediction of pathways contributes to understanding of alkaloid synthesis mechanisms and the application of graph based neural network models to similar problems in bioinformatics would therefore be beneficial.

Development of data-driven chemistry in chemistry and chemical engineering

Cheminformatics has been applied to various areas of chemistry: molecular design, materials design, organic synthesis design, structure elucidation and process control. Kimito Funatsu presented an overview of these applications during his research life. In pursuit of a desired function, a novel compound, material, or device is required. The first step in producing one is to decide what to make. Designing the molecule or material may involve modeling, inverse analysis, or data analysis. The next step is deciding how to make the product, and this may involve synthesis design or product prediction. A production process is then needed to make a commercial product. Reaction, separation, and refinement are carried out in the chemical plant, and, in order to produce the product with the desired property, quality control of the process is important. The fourth step is analysis, which may require structure determination.
Finally, after the products are provided to the public, methods for recycling and reuse are also required. These five units are important subjects in cheminformatics. Knowledge to support them is created from many kinds of data and information.

This knowledge has to be organized, by data modeling, and used for prediction and design. Structure-property (or activity) relationship models can be constructed, and candidate molecules or materials can be generated that satisfy the desired property, by “inverse analysis” (see Figure 13). The generation of candidate structures controlled by the model is the driving force for de novo design (in drug discovery), design of highly functional polymers (including monomer design), and catalyst design. Developing a structure generator is challenging; even for a molecular formula as simple as C₆H₆, there are 217 possible isomers.

In the first stage of development of new drugs, various lead compounds with high activity are required. To design such compounds, Kimito and co-workers have focused on chemical space defined by structural descriptors. New compounds close to areas where highly active compounds exist will show the same degree of activity. Visualization of chemical space is useful for understanding activity distribution in chemical space and determination of the target area for structure search by activity distribution. Structures in chemical space are described by many descriptors, giving rise to high dimensional chemical space. This is projected onto a 2D plane by generative topographic mapping. The activity is displayed as a heat map on this 2D plane. Thus target areas for structure search can be assigned.

Kimito’s team has developed a new de novo design system to search a target area. First, highly active compounds are manually selected as initial seeds. Then, the seeds are entered into the system, and structures slightly different from the seeds are generated and pooled. Next, seeds are selected from the new structure pool based on the distance from target coordinates on the map. Activity distribution and druglikeness can be visualized on the same map, and the target area selected by considering overlap. The initial de novo design system for exploring chemical space (DAECS) was modified to enable the user to select a target area to consider properties other than activity, and improve the diversity of the generated structures by visualizing the druglikeness distribution and the activity distribution, generating structures by substructure-based structural changes, including addition, deletion, and substitution of substructures, as well as the slight structural changes used in DAECS. Through case studies using ligand data for the human adrenergic alpha2A receptor and the human histamine H1 receptor, it has been shown that the modified DAECS can generate high diversity druglike structures, and the usefulness of the modification of the DAECS has been verified.
In the recent study, where the target protein was the histamine H1 receptor, the training data were 522 structures and pK\textsubscript{i} values selected from ChEMBL, and the descriptors were 142 fingerprints from PubChem. The training data for construction of the discriminant model were 1000 structures from BIOVIA Comprehensive Medicinal Chemistry (http://accelrys.co.jp/products/databases) and 1000 non-drug structures from the BIOVIA Available Chemicals Directory (http://accelrys.co.jp/products/databases). The modeling method was support vector machine. Kimito showed visualization with structure generation (Figure 14) and some generated structures (Figure 15).

Kimito next discussed polymer alloys, a class of polymer blends where addition of a second polymer is tailored to provide controlled morphology and thus specific performance characteristics. Polymer alloys can be produced by mixing, melting, and crystallizing a mixture of multiple polymers, then by molding, melting, and crystallizing the mixture. Data items include the properties of each component polymer, and the mixing conditions, molding conditions, and alloy properties. There can be 100-200 data items.

Kimito has also worked on the design of more efficient polymeric optical films (Figure 16) that manage the polarization of light. He explained the mechanism of polarizing transmittance and reflection, and its relationship to polymer orientation in machine direction (MD), and polymer orientation in transverse direction (TD).
His objective was to construct a quantitative model of the properties of light improved film and to design a more efficient film. He aimed to optimize process conditions such as extrusion and to achieve brightness (in cd/m²) ≥5400, MD transmittance ≥82%, and TD transmittance ≤20%. Object variables were brightness, MD transmittance, and TD transmittance. Explanatory variables were composition (percentages of polyethylene naphthalate, polyethylene terephthalate, and polystyrene), percentages of three compatibilizing agents, and process conditions (stretching temperature, extrusion machine ID (1 or 2), stretching magnification, and thickness). The number of samples was 26. The results of partial least square analysis were excellent: brightness $R^2 = 0.916$, $Q^2 = 0.682$, MD transmittance $R^2 = 0.977$, $Q^2 = 0.920$, TD transmittance $R^2 = 0.930$, $Q^2 = 0.746$.

Kimito’s final topic was process control. In operating chemical plants, operators have to monitor the operating condition of the plants and control process variables. So, process variables such as temperature, pressure, liquid level, and concentration of products need to be measured online, but none of them is easy to measure online because of technical difficulties, large measurement delays, high investment cost, and so on. In order to cope with this problem, soft sensors are widely used in chemical plants. Soft sensors are inferential models constructed between easy-to-measure variables, such as temperature and pressure, and variables that are difficult to measure online, such as concentration or property. By inputting temperature and pressure variables to soft sensor models, the soft sensor can estimate property and concentration variables online with high accuracy. Thus, the process operator can obtain, and use the predicted values for process control in real time.

Kimito returned to his initial theme of the steps in cheminformatics, the first step being what to make and the second being design. In the design step, a structure-property relationship model is constructed for inverse analysis, to generate candidate structures or materials. It is important to incorporate process parameters into the modeling step. How to make the material is considered at the same time. This is an important concept in materials design, because the property of materials is strongly affected by process conditions, even for the same starting materials. Eventually, a production process is needed to make a commercial product. Here process control, namely quality control of the product, is particularly important. In this step, the quality of the product is monitored by a soft sensor online, and the quality is controlled by operating process parameters. Simultaneous consideration of quality control can realize integrated treatment of materials design, examination of process conditions, and quality control. Kimito emphasizes this concept as process informatics.

Conclusion

Elsa Alvaro, chair of the ACS Division of Chemical Information, formally presented the Herman Skolnik Award to Kimito Funatsu at the end of the symposium.
References


(10) Sakanoue, K.; Motoda, M.; Sugimoto, M.; Sakaki, S. A Molecular Orbital Study on the

*Chemical Information Bulletin, 2019, 71* (4)
Chemical Information Bulletin, 2019, 71 (4)


(20) Schneider, P.; Schneider, G. De Novo Design at the Edge of Chaos. J. Med. Chem. 2016, 59 (9), 4077-4086.


(25) Schneider, G. De novo design - hop(p)ing against hope. Drug Discovery Today Technol. 2013, 10 (4), e453-e460.


(33) Schmuker, M.; Schneider, G. Processing and classification of chemical data inspired by insect
Chemical Information Bulletin, 2019, 71 (4)


Wendy A. Warr
Wendy Warr & Associates
wendy@warr.com
New Directions for CINF: Strategic Plan, 2020-2023

It's déjà vu all over again. Back in the spring of 2007, I was part of the CINF Strategic Planning Implementation Team that participated in a facilitated retreat to create a CINF strategic plan for 2007-2009. Now, as Chair of CINF for 2020, I am excited to unveil a new CINF vision, mission, and strategic plan for 2020-2023.

This strategic planning process was similar to the last one: a group of 12 CINF members and allies gathered for a strategic planning retreat to draft goals and strategies that would allow CINF to move forward for the next several years. This time, the group met at ACS Headquarters in Washington, DC, working over a day and a half (November 9-10, 2019) and facilitated by Carol Duane (a past chair of CINF) and Larry Krannich.

The team’s work was informed by data collected specifically for the strategic planning process, CINF membership data from ACS, the results of a CINF membership survey that was distributed a few months prior to the retreat, and a survey completed by retreat participants immediately before the retreat.

Attendees participated in a number of activities that allowed us to draft new vision and mission statements for CINF. Next, we worked to develop goals that would successfully address this new vision and mission, but we also considered the strengths, weaknesses, opportunities, and threats facing CINF. The group identified potential actionable strategies that would help CINF to achieve the goals successfully. We prioritized three strategies for each of the three goals, and the final retreat sessions had us breaking into one- or two-person groups to draft implementation plans, including a suggested team leader, potential team members, and a timeline of significant events, and to provide feedback on others’ plans. Those attending the retreat were asked to champion one of the strategies and to assist the eventual team leaders in understanding the context of their particular strategy within the broader strategic plan.

After the retreat, the language of the vision, mission, and goals was modified for clarity and consistency, and the resulting plan was presented to the CINF Executive Committee for a vote and was enthusiastically supported.

While the implementation teams still need to be charged with their work, and, therefore, timelines may be slightly adjusted, I wanted to share with you our new strategic plan:

**CINF Vision:** Better science through the power of chemical data, information, and knowledge.

**CINF Mission:** Prepare and empower the scientific community to create, analyze, organize, and disseminate chemical information and data.

**Goal 1:** Establish and implement a strategic framework and infrastructure for CINF organizational sustainability.

- Strategy 1: Implement a new fundraising and finance model for all divisional activities by the beginning of 2021.
- Strategy 2: Submit to the CINF Executive Committee a proposal related to a comprehensive incentive, recognition, and awards program by the fall national meeting in 2020.
- Strategy 3: Develop a mentorship and personal development program for the CINF leadership pipeline by the fall national meeting in 2020.
Goal 2: Develop quality, innovative activities and a supportive community that empower, engage, and serve our members.

- Strategy 1: Launch a re-designed CINF website by the end of 2020.
- Strategy 2: Plan, research, and execute three training areas and their content for librarian-ship, cheminformatics, data science (chemical) by the end of 2020.
- Strategy 3: Create a new committee chair position leading communications and social media by the start of 2021.

Goal 3: Expand multi-disciplinary collaboration and outreach within and beyond ACS in order to connect globally and share knowledge.

- Strategy 1: Correlate membership data to establish liaisons with other divisions by the end of 2021.
- Strategy 2: Establish a liaison program between CINF and non-ACS groups through a divisional survey by the end of 2021.
- Strategy 3: Identify international organizations interested and willing to host joint activities with CINF by the end of 2022.

Overall, the first goal focuses on making CINF more sustainable and organizationally sound as a division. The second goal is meant to enhance the value that the division offers to our members, while the third goal encourages CINF to collaborate with other organizations, both internal and external to ACS, while also paying attention to our international members and partners.

Special thanks to Elsa Alvaro who assisted greatly in the planning of the retreat and communicated with ACS staff about logistics for the event. At ACS, we were assisted by Cassaundra Evans and Melissa Paluch.

Please be on the lookout from CINF about how you can get involved with the implementation teams: how you might join one of them, how you can provide feedback on their work, or how you can take part in any new offerings created by these teams. I also welcome your feedback on the plan and any ideas you might have for any (or all) of the strategies.

Here is to a prosperous 2020.

Jeremy Garritano
University of Virginia
jg9jh@virginia.edu
CINF Strategic Planning Retreat Participants

November 9-10, 2019
Washington, DC

Back Row (left to right): Larry Krannich (facilitator), Michael Qiu, Donna Wrublewski, Emilio Esposito, Alex Williams, Evan Bolton, Rajarshi Guha, Vin Scalfani, Stuart Chalk.

Front row (left to right): Sue Cardinal, Annie Zeidman-Karpinski, Jeremy Garritano, Judith Curran, Carol Duane (facilitator)
CINF Luncheon: The End of a Tradition

The last CINF Luncheon was organized at the 2019 spring ACS national meeting in Orlando. The event banner on the left was from this meeting (photos from Orlando CINF Luncheon, [https://www.flickr.com/photos/cinf/albums/72157704683251272](https://www.flickr.com/photos/cinf/albums/72157704683251272)). The Luncheon was not held during the fall meeting in San Diego and has not been planned for the two 2020 meetings.

The Luncheon existed for over 75 years and was the earliest CINF social event. In “50 Years of Chemical Information in the American Chemical Society 1943-1993.” ([http://web.stanford.edu/group/swain/cinf/50years/html_index.html](http://web.stanford.edu/group/swain/cinf/50years/html_index.html)), Val Metanomski diligently recorded, “At almost every ACS National Meeting, the Division members enjoyed having a Divisional luncheon, usually on a Tuesday... The tradition was started with the Chemical Literature Group in 1943, long before the Division was formed in 1948.”

I quoted the above excerpt in my earlier article about the recent history of the CINF Luncheon, published in the summer 2017 issue of the Chemical Information Bulletin ([https://bulletin.acscinf.org/node/956](https://bulletin.acscinf.org/node/956)). The warning signs of the event’s uncertain future grew consistently over the past decade. The division has been trying to retain the Luncheon through multiple efforts: finding a generous sponsor; reducing event expenses by going from a full service to a buffet style lunch; choosing a less expensive, official ACS venue, such as the convention center instead of a hotel; and increasing the individual ticket cost. Yet, despite all attempts, the Luncheon’s days were numbered. In 2017, the division had to deal with the highest costs recorded for this event, reaching almost $10,000 for the spring meeting in San Francisco and almost at $9,000 for the fall meeting in Washington. Such high costs were detrimental to the division’s priorities, for receptions were better venues for social networking among CINF members and friends. Thanks for the generous support of the CINF Luncheon should be extended to the Royal Society of Chemistry, who have been sponsoring the event exclusively for the past eight years and have responded with additional contributions to offset costs in 2017 and 2018.

According to the ACS national meeting registration data, the overall attendance does not correlate well with the ticket sales for the CINF Luncheon. In addition to sales via ACS registrations and CINF on-site transactions, the division has always offered complimentary tickets to the division’s award winners and keynote speakers. Indeed, the actual attendance at the CINF Luncheons has increased, reaching up to 80-90 people, although only 55-65 tickets were sold in advance. The latest procedural implementation by ACS for the spring meeting in Orlando added an additional obstacle for the Luncheon by requiring the division to pay in advance for all tickets. Thus, the low number of 44 tickets sold was not just the lowest overall sales figure, but it also included several tickets purchased by CINF. Taking into consideration the higher-than-expected attendance for the Orlando meeting, such small sale was disappointing, and the division did not want to keep increasing the individual ticket price. The ticket price increase for the 2018 meeting in Boston was for a special location and anticipated high on-site participation by the chemical information community.
In advance of the 2019 San Diego meeting, the CINF executive committee estimated a high cost for the Luncheon and voted against holding it. The decision was unanimous, and nobody on the executive committee expressed strong feelings for sustaining this traditional event in its current mode. There were a few discussions of ideas for alternative, less expensive options for that time slot, but for the time being, focus is shifting to the receptions. Fundraising is an arduous task. Fundraising for the 2019 Orlando meeting yielded extremely low contributions. A few years ago, the division allocated a small budget for a reception at each of the two national meetings as a preventive measure, in the event that they could not secure sufficient external funds, but CINF relies substantially on the generous support of its sponsors for all of its social events. The 2019 San Diego meeting saw very successful fundraising. Obviously, in San Diego, CINF joined the majority of the divisions in not organizing a luncheon, a type of social function that is more extravagant than other events. Based on the San Diego meeting program, the organizers of luncheons included only two divisions, two committees, and the core student programming group, which held a lecture with a luncheon.

- Division of Colloid and Surface Chemistry Luncheon $45.00 per ticket
- Division of Chemistry and the Law Networking Luncheon $40.00 per ticket
- Committee on Minority Affairs Luncheon $50.00 per ticket
- Women Chemists Committee Luncheon $50.00 per ticket
- Eminent Scientist Lecture & Luncheon (Core Student Programming) $35.00 per ticket

In conclusion, I can speculate that the modern lifestyle and food culture trends have influenced our choices. As a result, the traditional dining rituals are becoming excessively expensive, both for CINF members and for our sponsors.

Acknowledgements:

I would like to express special thanks to Brenda Philpot for providing data from ACS national meetings, Stuart Chalk for commenting about CINF finances, and Wendy Warr for managing the CINF flickr website.

Appendix: Tables 1 and 2 update the charts in my previous article, which ends with spring 2017.

Table 1: CINF Luncheon Speakers 2017-2019

<table>
<thead>
<tr>
<th>Meeting</th>
<th>Speaker</th>
<th>Presentation Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spring 2017</td>
<td>Andrew Leach</td>
<td>Molecules, data, and models</td>
</tr>
<tr>
<td>San Francisco</td>
<td>EMBL-EBI</td>
<td></td>
</tr>
<tr>
<td>Fall 2017</td>
<td>Kim Pruitt</td>
<td>Managing terabytes petabytes of data to support connecting data to knowledge</td>
</tr>
<tr>
<td>Washington</td>
<td>NCBI</td>
<td></td>
</tr>
<tr>
<td>Spring 2018</td>
<td>John Pardue</td>
<td>Chemical releases during disasters: lessons from Katrina to Harvey</td>
</tr>
<tr>
<td>New Orleans</td>
<td>Louisiana State University</td>
<td></td>
</tr>
<tr>
<td>Fall 2018</td>
<td>Alex Clark</td>
<td>Leveling up chemical information for the era of big data</td>
</tr>
<tr>
<td>Boston</td>
<td>Collaborative Drug Discovery</td>
<td></td>
</tr>
<tr>
<td>Spring 2019</td>
<td>Wendy Warr</td>
<td>How to right a scientific paper</td>
</tr>
<tr>
<td>Orlando</td>
<td>Wendy Warr &amp; Associates</td>
<td></td>
</tr>
</tbody>
</table>
Table 2: CINF Luncheon Expenses 2017-2019

<table>
<thead>
<tr>
<th>Meeting Location</th>
<th>Meeting Attendance</th>
<th>Ticket Price</th>
<th>Number of tickets sold via ACS</th>
<th>Amount received via ACS registrations</th>
<th>Amount received via CINF sales on</th>
<th>Amount received from the sponsor</th>
<th>Amount billed by ACS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spring 2017 San Francisco</td>
<td>18,917</td>
<td>$30</td>
<td>52</td>
<td>$1560</td>
<td>$240</td>
<td>$4000</td>
<td>$9592</td>
</tr>
<tr>
<td>Fall 2017 Washington</td>
<td>12,944</td>
<td>$30</td>
<td>57</td>
<td>$1710</td>
<td>$150</td>
<td>$5000</td>
<td>$8701</td>
</tr>
<tr>
<td>Spring 2018 New Orleans</td>
<td>16,752</td>
<td>$30</td>
<td>65</td>
<td>$1950</td>
<td>$90</td>
<td>$4000</td>
<td>$4298</td>
</tr>
<tr>
<td>Fall 2018 Boston</td>
<td>14,463</td>
<td>$35</td>
<td>64</td>
<td>$2240</td>
<td>$175</td>
<td>$4500</td>
<td>$6493</td>
</tr>
<tr>
<td>Spring 2019 Orlando</td>
<td>15,754</td>
<td>$30</td>
<td>44</td>
<td>$1320</td>
<td>$0</td>
<td>$3500</td>
<td>$3960</td>
</tr>
</tbody>
</table>

Svetlana Korolev
University of Wisconsin-Milwaukee
skorolev@uwm.edu
Book Review


This manual of good practice is intended for those concerned with the evaluation of the value and impact of research: administrators, scholars, regulators, politicians, and other members of the public. This book will broaden the scope (and hopefully the quality) of the evaluation of research, an area previously dominated by specialists (scientometricians) and analysts. The book consists of four chapters in Q&A format, a section on further reading, and an index.

Chapter 1, “Basics”, covers why one should measure research, the historical and theoretical measurement methods (largely bibliometrics), the relationship between indicators and concepts, and the available data sources. Standards for the data sources have been developed by two international organizations, the Organisation for Economic Co-operation and Development (OECD) and the United Nations Educational, Scientific and Cultural Organization (UNESCO). Bibliographic databases of the published literature are important sources of data, but few comprehensive databases of books exist at this time.

Chapter 2 describes the data. Citation indexes are covered in general, including Web of Science and the Science Citation Index, Scopus, and Google Scholar. The chapter also covers the differences between the sources and their historical performance, as well as cultural biases and definitions of disciplines.

Chapter 3 is the longest chapter, and it discusses the indicators. The definition of authorship and resulting measurement is very important. Related to this are the definition and measurement of research production and collaboration, and the interdisciplinarity of research. Differences between references and citations are discussed, as is the variability of citation rates (with several graphs). The chapter discusses reasons both for citing and for not citing research (for example, over a given period of time, half of the available papers are never cited). Methods of counting citations and their concentration are discussed, and obsolescence and the controversial issues of self-citation and self-referencing are also mentioned.

Journal impact factors, Eigenfactor Scores, Source Normalized Impact Factors, SCImago journal rankings, and CiteScores are all discussed. The h-index and variants are defined, but the reviewer would have appreciated more discussion of the h-index and variants; for more detail, a reader must go elsewhere. The authors also describe altmetrics, a concept that has spawned a large number of organizations providing such data and analysis. The measurement of research funding and the relationship between science indicators and peer review are both described.

Chapter 4 is titled, “The Big Picture”. The big questions addressed involve the people or entities who control the measurement of research, the identification of the stakeholders and their responsibilities, the adverse effects of the measurement of research (including effects on individuals, incentives, etc.), and the future of the field. The last depends on application (and misapplication) of the methods. A quote from the last paragraph of the chapter is a good summary, “Research evaluators would do well to heed the mantra of the medical profession: First do no harm.” Hopefully, use of this book will help.

This book should be available to and used by individuals directly affected by measurements of research and by people who serve clients who are affected.

Robert E. Buntrock
Buntrock Associates, Orono, ME
buntrock16@roadrunner.com
Committee Reports

Report from the Council Meeting Held on August 28, 2019

The Council of the American Chemical Society met in San Diego, CA, on Wednesday, August 28, 2019, from 8:00 a.m. until approximately 11:30 a.m. in the Sapphire Ballroom of the Hilton San Diego Bayfront Hotel. Below are the highlights of the meeting.

Council Actions:

Elections for the Committees of Council
Council Policy Committee: Anne M. Gaffney, Lydia E. M. Hines, Will E. Lynch, and Sally B. Peters were elected for three-year terms (2020-2022), and Dee Ann Casteel was elected for a one-year term (2020).

Committee on Nominations and Elections: Michelle V. Buchannan, Charles E. Cannon, Alan A. Hazari, Amber S. Hinkle, and Thomas H. Lane were elected for three-year terms (2020-2022).

Committee on Committees: Lisa M. Balbes, D. Richard Cobb, Emilio X. Esposito, Jason E. Ritchie, and Stephanie J. Watson were elected for three-year terms (2020-2022).

Other Council Actions
On the recommendation of the Committee on Committees, and with the concurrence of the Council Policy Committee, the council approved the continuation of the Committees on International Activities and on Professional Training, contingent on approval by the Board of Directors.

On the recommendation of the Committee on International Activities, and with the concurrence of the Council Policy Committee, the council approved the creation of an ACS International Chemical Sciences Chapter in the Republic of Georgia, contingent on approval by the Board of Directors.

On the recommendation of the Committee on Nominations and Elections, the council voted that the Pittsburgh Local Section be transferred from District II to District III in order to bring District III’s member population into compliance with bylaw requirements.

Resolutions
The council passed the following resolutions: a resolution in memory of former Executive Director John Kistler Crum; a resolution in memory of other deceased Councilors; a resolution in recognition and celebration of the 100th birthday of Gerald Meyer and his 80 years of service to the society as an ACS member; a resolution in gratitude to the officers and members of the San Diego Local Section (host section for the 258th National Meeting), to the divisional program chairs and symposium organizers, and to ACS staff; and a resolution acknowledging Bonnie A. Charpentier’s service as ACS president and presiding officer of the council.

Highlights from Committee Reports:

Nominations and Elections (N&E)
The Committee on Nominations and Elections solicits councilors’ recommendations of qualified individuals for the positions of President-Elect and Directors. Suggestions can be sent to nomelect@acs.org.
Ballots for the 2019 fall national election will be distributed starting on September 30, with a voting deadline four weeks later on October 25. N&E encourages all ACS members to vote for president-elect and on the constitutional amendment. Election information may be viewed at https://www.acs.org/elections.

CINF councilors recommended that CINF members vote to adopt the constitutional amendment. We sincerely believe it is in the best interest of the society and our division. These changes will provide more flexibility and responsiveness in ACS governance.

**Budget and Finance (B&F)**

The society’s 2019 financial performance through July 31 yielded a net from operations of $30.1 million. This is $10 million favorable to the approved budget, and $1.7 million less than it was over the same period in 2018. Total revenues are on budget at $338 million. Total expenses are $308 million, which is $10 million favorable to budget.

The board reviewed committee proposals for funding for the ChemIDP, the International Student Chapters Programs, and Green and Sustainable Chemistry Education Resources pilot in the 2020 proposed budget.

The society expects to end the year in compliance with each of the five board-established financial guidelines. Additional information can be found at https://www.acs.org; at the bottom of the page, click “About ACS”, and then click “Financial”.

**San Diego Meeting Attendance**

The theme of the 258th ACS National Meeting was “Chemistry and Water”. As of Tuesday evening, August 27, attendance was:

| Attendees | 7488 |
| Students  | 3095 |
| Exhibitors| 995  |
| Expo only | 430  |
| Guest     | 401  |
| **Total** | **12,409** |

Fall national meeting attendance since 2004 is as follows:

<table>
<thead>
<tr>
<th>Year</th>
<th>Location</th>
<th>Attendance</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td>Philadelphia, PA</td>
<td>14,025</td>
</tr>
<tr>
<td>2005</td>
<td>Washington, DC</td>
<td>13,148</td>
</tr>
<tr>
<td>2006</td>
<td>San Francisco, CA</td>
<td>15,714</td>
</tr>
<tr>
<td>2007</td>
<td>Boston, MA</td>
<td>15,554</td>
</tr>
<tr>
<td>2008</td>
<td>Philadelphia, PA</td>
<td>13,805</td>
</tr>
<tr>
<td>2009</td>
<td>Washington, DC</td>
<td>14,129</td>
</tr>
<tr>
<td>2010</td>
<td>Boston, MA</td>
<td>14,151</td>
</tr>
<tr>
<td>2011</td>
<td>Denver, CO</td>
<td>10,076</td>
</tr>
<tr>
<td>2012</td>
<td>Philadelphia, PA</td>
<td>13,251</td>
</tr>
<tr>
<td>2013</td>
<td>Indianapolis, IN</td>
<td>10,840</td>
</tr>
<tr>
<td>2014</td>
<td>San Francisco, CA</td>
<td>15,761</td>
</tr>
<tr>
<td>2015</td>
<td>Boston, MA</td>
<td>13,888</td>
</tr>
<tr>
<td>2016</td>
<td>Philadelphia, PA</td>
<td>12,800</td>
</tr>
<tr>
<td>2017</td>
<td>Washington, DC</td>
<td>12,904</td>
</tr>
<tr>
<td>2018</td>
<td>Boston, MA</td>
<td>14,235</td>
</tr>
<tr>
<td>2019</td>
<td>San Diego</td>
<td>12,409</td>
</tr>
</tbody>
</table>

The attendance at the San Diego meeting is lower than expected, but is counterbalanced by a strong turnout in Orlando for the spring 2019 meeting.
The Career Navigator live program was well attended. There were 31 participating employers, 81 available positions, 239 job seeker profiles, 270 lightning talk attendees, 118 Career Pathway registrations, and 555 career consultant interactions.

**Petitions to Amend the Constitution and Bylaws**

There were no petitions up for action in San Diego.

New petitions to amend the constitution or bylaws must be received by the Executive Director & CEO no later than December 4 to be included in the Council agenda for consideration at the spring 2020 meeting in Philadelphia. Contact the Committee on Constitution and Bylaws at bylaws@acs.org with any questions or requests for information.

**Actions of the Board of Directors:**

**The Board’s Executive Session**

At this meeting, the ACS Board of Directors focused on a number of key strategic issues and took several related actions.

**The Board’s Committees**

The board discussed reports from its committees on Budget and Finance, Professional and Member Relations, Strategic Planning, Executive Compensation, and Pensions and Investments, as well as the ACS Governing Board for Publishing, the Task Force on the Future of Meetings, and the Society Programs Globalization Board liaison. In particular:

- On the recommendation of the Committee on Professional and Member Relations, the board voted to approve the society’s nominees for the 2020 National Science Board Public Service Award and the 2020 Tang Prize in Biopharmaceutical Science.

- On the recommendation of the Joint Board-Council Committee on Publications and an Editor Selection Committee, the board voted to approve the appointment and reappointment of editors-in-chief for ACS journals. Their names will be announced once they have been notified and practical arrangements for their service to ACS have been finalized.

- On the recommendation of the Society Committee on Budget and Finance, the board voted to set the advance member registration fee for national meetings held in 2020 at $505 (this amount is based on the 2019 fee adjusted for inflation) and to approve several program funding requests. While, in recent years, an additional escalator has been applied, meetings are breaking even over a five-year period, so, no additional amount was recommended for 2020.

**Executive Director and CEO Report**

The Executive Director and CEO and his direct reports provided updates to the board on the activities of Chemical Abstracts Service (CAS) and the ACS Publications divisions. He offered updates on issues relating to the ACS Core Value of Diversity, Inclusion, and Respect; the current status of society membership; ACS financials; initiatives associated with the International Year of the Periodic Table; and upcoming events and activities. As part of his report, he invited the Executive Vice President for Scientific Advancement to lead an informal discussion on key issues for that division.

**Other Society Business**

The board heard reports from members of the presidential succession on their current activities as well as those planned for 2020, particularly the presidential symposia and endorsed symposia for the San Diego meeting.
The board liaison for globalization provided a summary of a recent board survey and received additional feedback from the board on the globalization vision for society programs. The goal here is encouragement and expansion, where appropriate, of existing successful international activities and initiatives, as well as evaluating current products and programs; exploring additional options and opportunities; and advising the board on the assembly of a coherent and balanced program portfolio appropriate to the globalized ACS of the twenty-first century.

The board received a preview of recommendations for process improvements focused on streamlining the ACS strategic planning process from the Committee on Strategic Planning, including an adjustment of the cycle time, inclusion of an analysis of professional association market dynamics, and concentration on only the highest-value elements for the Strategic Plan.

The board received its customary extensive briefing from its Committee on Executive Compensation. The compensation of the society’s executive staff continues to receive regular review by the board.

The board received a status update from the Task Force on the Future of Meetings. The task force has been charged with performing a “deep dive” on the current portfolio of ACS meetings and conferences; identifying current offerings; evaluating governance and staff support structures, revenue streams, financial targets, and business models; and recommending actions that will ensure the sustainability and future relevance of that portfolio.

The board also held a discussion with members of the executive committee and leadership staff of the Society for the Advancement of Chicanos/Hispanics and Native Americans in Science (SACNAS) on the vision, mission, and structure of SACNAS. Various components of the recently signed Chemistry Enterprise Partnership agreement were discussed, emphasizing the shared overarching objective of promoting and achieving diversity in STEM. The board agreed to the appointment of a joint ACS-SACNAS task force charged with developing additional short-term and long-term goals, and evaluating continued partnership potential.

**The Board’s Regular Session**

The Board held a well-attended interactive regular session on Sunday, August 25, that featured the work done by the Task Force on the Next Generation Leadership Program portfolio as well as a celebration of the tenth anniversary of the ACS Leadership Development System.

**Supplemental information**

2019 ACS Officers
Luis Echegoyen, President-Elect
Bonnie A. Charpentier, President
Peter K. Dorhout, Immediate Past President

ACS Offices
Office of Secretary & General Counsel

Committees
Budget and Finance
Chemical Safety
Chemistry & Public Affairs
Committee on Committees
Co.nstitution and Bylaws
Economic and Professional Affairs
Environmental Improvement
Ethics
Local Section Activities
Minority Affairs

Nominations and Elections
Patents and Related Matters
Professional Training
Technician Affairs

Women Chemists
Younger Chemists

Resources on the Web
Yellow Book
Highlights of ACS Achievements
Governing Documents
Chemists with Disabilities
Community Activities Outreach
Project SEED free webinar
CPRC PR Makeover

Submitted by the CINF Councilors

Svetlana Korolev
CINF Councilor
skorolev@uwm.edu
Bonnie Lawlor
CINF Councilor
chescot@aol.com
Andrea Twiss-Brooks
CINF Councilor
atbrooks@uchicago.edu
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 2019 National Meeting in San Diego, CA, 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**
- ACS Publications
- Schrodinger

**Silver**
- Google
- Wiley

**Bronze**
- Chemical Abstracts Service
- Elsevier Reaxys
- Thieme Chemistry

**Contributor**
- Journal of Chemical Information & Modeling (ACS Publications)
- Bio-Rad Laboratories
- InfoChem

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:


Please feel free to contact me if you would like more information about supporting the ACS Division of Chemical Information.

Tina Qin  
Chair, Fundraising Committee 2020  
Email: Sponsorship@acscinf.org

The ACS CINF Division is a non-profit tax-exempt organization with taxpayer ID no. 52-6054220.
Bio-Rad Laboratories

Featured Academic Resources from Bio-Rad

KnowItAll U - World’s Largest Spectral Database

KnowItAll U offers every student and faculty member unlimited access to over 2 million spectra, including Infrared (IR), Raman, NMR (CNMR, HNMR, XNMR), UV-Vis, and Mass Spectra (MS).

KnowItAll U Spectral Library subscriptions provide the most affordable option for campus-wide access to the largest and highest quality collection of spectra available today.

Learn More: http://www.knowitall.com
Request Trial: http://www.knowitall.com/contactus

Choose access to the complete KnowItAll U collection or spectra from a specific technique

<table>
<thead>
<tr>
<th>Product Code</th>
<th>Library Name</th>
<th>Spectra</th>
</tr>
</thead>
<tbody>
<tr>
<td>877010</td>
<td>KnowItAll U (all techniques)</td>
<td>2,074,000</td>
</tr>
<tr>
<td>891050</td>
<td>KnowItAll U - IR</td>
<td>264,000</td>
</tr>
<tr>
<td>892050</td>
<td>KnowItAll U - MS</td>
<td>860,000</td>
</tr>
<tr>
<td>895050</td>
<td>KnowItAll U - NMR</td>
<td>895,000</td>
</tr>
<tr>
<td>894050</td>
<td>KnowItAll U - Raman</td>
<td>25,000</td>
</tr>
<tr>
<td>896050</td>
<td>KnowItAll U - UV-Vis</td>
<td>30,000</td>
</tr>
</tbody>
</table>

KnowItAll Academic Edition - Free Software for Chemistry Students & Professors

The KnowItAll Academic Edition is a FREE chemistry software package for the academic community. Tools include ChemWindow chemical structure drawing software / reporting & publishing tools, along with spectral analysis tools.

Download Now: www.knowitall.com/academic
Sponsor Announcements

Thieme Chemistry

Organic Materials: New Open Access Journal

The Thieme Group is happy to introduce *Organic Materials*, an Open Access journal focusing on the interdisciplinary field of chemistry and materials. Its scope includes synthesis of functional organic materials, studies of structure–property relationships, and applications of organic materials in a broader sense. Experimental and theoretical studies are covered.

*Organic Materials* offers a rigorous peer review process and the **Pay-What-You-Want pricing model**. The authors decide how much, or how little, they would like to pay for the article processing charges!


The first manuscripts by renowned chemists can be found here:

2020 CINF Officers

Chair
Jeremy Garritano
University of Virginia
jg9jh@virginia.edu

Chair-Elect
Donna Wrublewski
Caltech Library
dtwrub@caltech.edu

Past-Chair
Elsa Alvaro
Northwestern University
elsa.alvaro@northwestern.edu

Secretary
Nick Ruhs
Florida State University
nruhs@fsu.edu

Treasurer
Stuart Chalk
University of North Florida
schalk@unf.edu

CINF Councilors
Bonnie Lawlor
chescot@aol.com

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

CINF Alternate Councilors
Rachelle Bienstock
RJB Computational Modeling LLC
rachelleb1@gmail.com

Charles Huber
University of California, Santa Barbara
huber@library.ucsb.edu

Archivist/Historian
Bonnie Lawlor
chescot@aol.com

Awards Committee Chair
Rajarshi Guha
Vertex Pharmaceuticals
rajarshi.guha@gmail.com

Careers Committee Chair
Neelam Bharti
Carnegie Mellon University Libraries
nbharti@andrew.cmu.edu

Communications and Publications Committee Chair
Graham Douglas (Interim)
communications@acscinf.org

Procedures Chair
Bonnie Lawlor
See Councilor

Education Committee Chair
Grace Baysinger
Stanford University
graceb@stanford.edu

Finance Committee Chair
Stuart Chalk
University of North Florida
schalk@unf.edu

Fundraising Committee Chair
Tina Qin
Harvard University
qinnamsu@gmail.com

Membership Committee Chair
Donna Wrublewski (interim)
Caltech Library
dtwrub@caltech.edu
2020 CINF Officers

Nominating Committee Chair
Elsa Alvaro
Northwestern University
elsa.alvaro@northwestern.edu

2019–2020 Program Committee Chair
Sue Cardinal
University of Rochester
scardinal@library.rochester.edu

2020–2021 Program Committee Chair
Ye Li
MIT
yel@mit.edu

Chemical Information Bulletin Editor, Spring
Teri Vogel and Judith Currano (interim)
See CIB Editors, Fall and Winter

Chemical Information Bulletin Editor, Summer
David Shobe
Patent Information Agent
avidshobe@yahoo.com

Chemical Information Bulletin Editor, Fall
Teri Vogel
UC San Diego Library
tmvogel@ucsd.edu

Chemical Information Bulletin Editor, Winter
Judith Currano
University of Pennsylvania
currano@pobox.upenn.edu

Webmasters
Rachelle Bienstock
RJB Computational Modeling LLC
rachelleb1@gmail.com
Stuart Chalk
University of North Florida
schalk@unf.edu

Articles and Features
Jeremy Garritano
Svetlana Korolev
Wendy Warr

Symposium Reports
Susan Cardinal
Wendy Warr

Committee Reports
Rajarshi Guha
Svetlana Korolev
Bonnie Lawlor
Andrea Twiss-Brooks

Book Reviewer
Robert E. Buntrock

Sponsor Information
Graham Douglas

Production
Judith N. Currano
Bonnie Lawlor
David Shobe
Wendy Warr