RSC INTEREST GROUP
CHEMICAL INFORMATION AND
COMPUTER APPLICATIONS GROUP

NEWSLETTER summer 2021

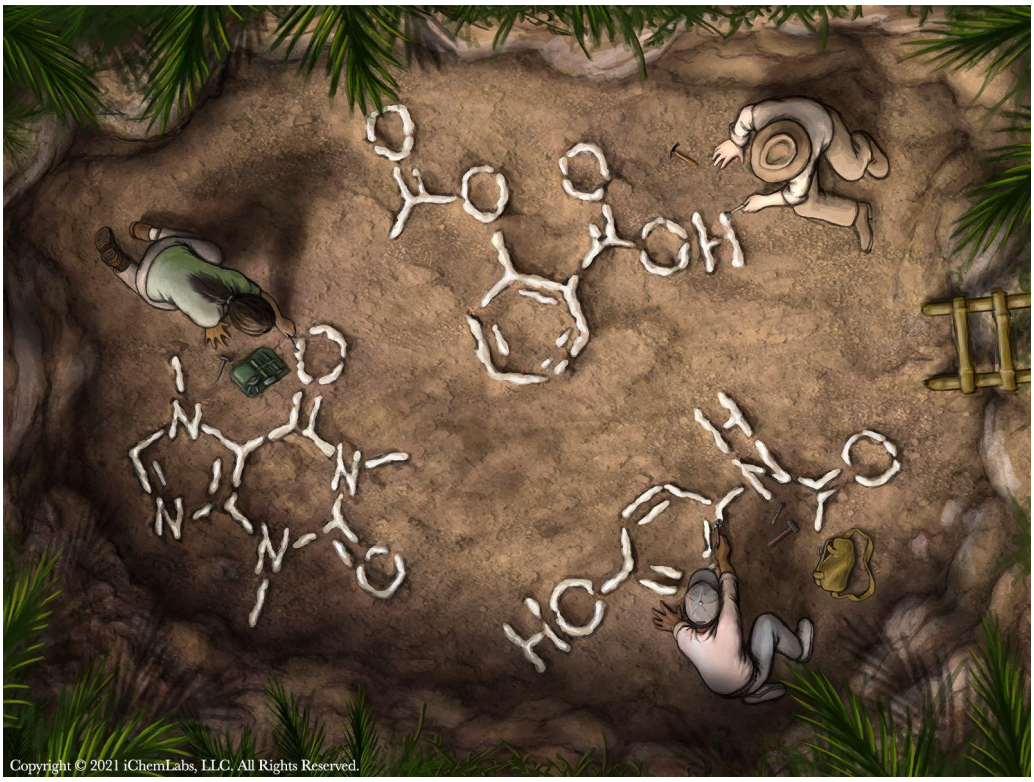


Image: "*Chemical Data Recovery 1: Embedded Chemical Data*" (see page 23)

CICAG aims to keep its members abreast of the latest activities, services, and developments in all aspects of chemical information, from generation through to archiving, and in the computer applications used in this rapidly changing area through meetings, newsletters and professional networking.

Chemical Information & Computer Applications Group Websites:

http://www.rsccicag.org
http://www.rsc.org/CICAG

YouTube https://www.youtube.com/c/RSCCICAG

LinkedIn http://www.linkedin.com/groups?gid=1989945

https://Twitter.com/RSC_CICAG

QR Code

Table of Contents

Contributions to the CICAG Newsletter are welcome from all sources - please send to the Newsletter Editor:

Stuart Newbold, FRSC, email: stuart@psandim.com

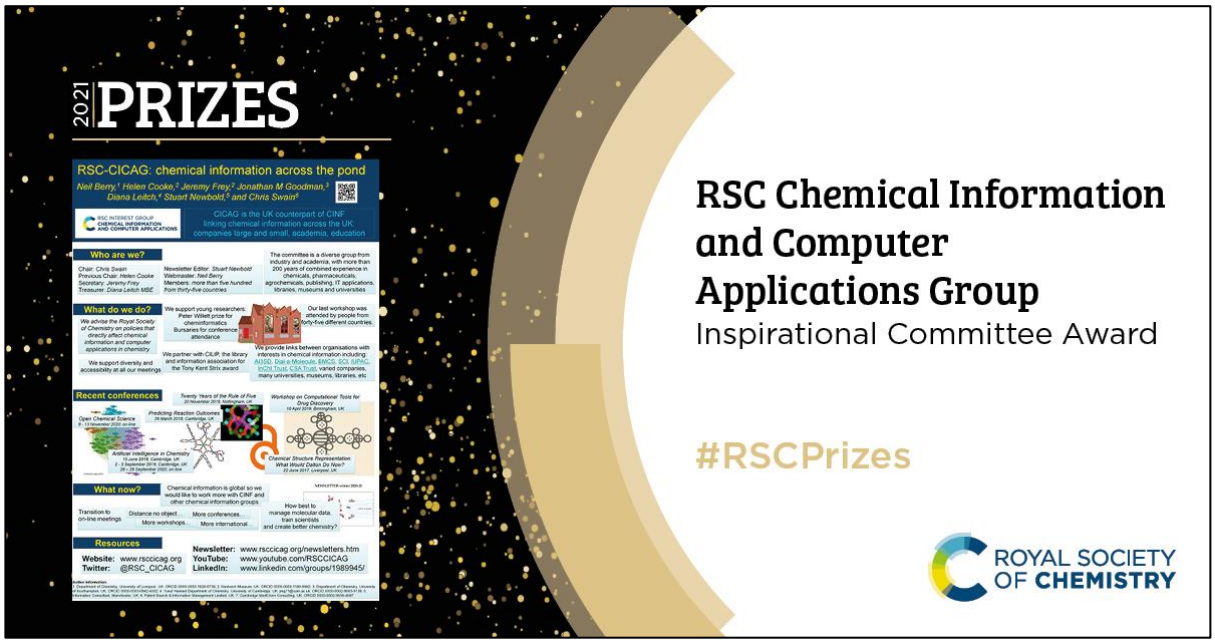
Chemical Information & Computer Applications Group Chair's Report

Contribution from RSC CICAG Chair Dr Chris Swain, email: swain@mac.com

The current COVID pandemic continues to cast a shadow over the world and our thoughts are with all those who have been affected.

Social media is becoming an increasingly active way for communicating with members during lockdown with Twitter having over 1250 followers, and LinkedIn now with 470 followers, gaining popularity. These channels provide a means of highlighting events including those from CICAG, plus other events that might be of interest. Our YouTube channel now contains video presentations from previous meetings and also all the Open-Source Tools for chemistry workshops, some of which have been viewed nearly 1000 times. There is clearly an unmet need for this type of workshop, and we have arranged monthly workshops until the end of the year. We are looking for potential suggestions for 2022.

The workshops arose following a 5-day meeting in November 2020 organised by RSC CICAG entitled Open Chemical Science. The online event proved to be enormously popular with attendees from 45 different countries. The challenges involved in converting what was planned as a physical event into a 5-day virtual event with 3 intertwined strands was recognised by the RSC with the award of the "2021 Inspirational Committee Award".



Our own CICAG website is often updated and we would be very interested to hear suggestions for additional content for all channels.

CICAG has collaborated with the AI3SD Network+ to create the AI 4 Proteins Seminar Series 2021 which culminated in the AI3SD & RSC-CICAG Protein Structure Prediction Symposium on 16-17 June. This 2-day event brought together many experts to discuss and debate the current challenges in predicting protein structure.

On June 7th the first meeting of underrepresented groups in Computational Chemistry was held entitled, "*Being #CompChemURG: How Diversity Enriches Us*". CICAG was delighted to support this event and there are more details in the newsletter (see page 16 for the meeting report).

This newsletter also includes contributions from Kurt Zielenbach, Christine Zardecki, Jürgen Jost & Guillermo Restrepo, Hannah Bruce Macdonald, Gianna Manchester, Kevin Theisen, Yvonne Pope and Samantha Kanza. CICAG is extremely grateful to receive these excellent pieces from external contributors.

Once again, I'd like to invite contributions that would be of interest to the CICAG community.

Sadly, two committee members Nathalie Batoux and Yvonne Pope are stepping down from the committee and I'd like extend my thanks for all their efforts. They will be missed. This does mean that we will seeking new additions to the committee in due course.

Whilst RSC members can join interest groups free of charge, in practice many members do not take up this opportunity. You can make a request to join a group via email (membership@rsc.org) or telephone (01223 432141).

CICAG Planned and Proposed Future Meetings

The table below provides a summary of CICAG's planned and proposed future scientific and educational meetings. For more information, please contact CICAG's Chair, Dr Chris Swain.

Meeting	Date	Location	Further Information
Big Data	TBD	Virtual/TBD	Proposed joint meeting with the SCI)
Open-source Software Workshops	Q3/4 2021	Virtual events	An ongoing series of workshops
4th Artificial Intelligence in Chemistry Meeting	27-28 Sep 2021	Virtual event	Joint event from RSC-CICAG and RSC-BMCS division
Chemical Information during Covid	TBD	Virtual/TBD	Details to follow
Python for Chemists	TBD	Virtual/TBD	Details to follow
Storing Chemical Data, why use a UDM?	TBD	Virtual/TBD	Details to follow

Open-Source Software Workshops

Contribution from RSC CICAG Chair Dr Chris Swain, email: swain@mac.com

All scientists working Chemistry need software tools for accessing, handling and storing chemical information, or performing molecular modelling and computational chemistry. There is now a wealth of open-source tools to help in these activities, however many are not as well-known as commercial offerings.

These workshops offer a unique opportunity for attendees to try out a range of open-source software packages for themselves with expert tuition in different aspects of chemistry. Attendees will be able to choose from sessions covering accessing online resources; data processing and visualisation; ligand and structure-based design, or computational chemistry. All software and training materials required for the workshop will be provided for attendees to install and run on their own computers.

Future meetings in 2021
Registration for these free workshops is now open. Our aim is to have a once a month workshop, usually the third week of each month on Thursday (or within a day or two).

22nd July 2-4 pm (UK time)
Charlotte Deane will host a workshop on the OPIG (Oxford Protein Informatics Group) antibody modelling tools. The OPIG Antibody Suite V2.0 is a collection of informatics tools and databases, including CoV-AbDab: the coronavirus antibody database.

19th August 2-4 pm (UK time)
Pat Walters (Relay Pharm and Practical Cheminformatics) provides an introduction to Cheminformatics, data analysis and machine learning. A hands-on workshop on building and validating ML models.

- Initial exploratory data analysis
- ML model building
- Model evaluation
- Making predictions on a larger dataset

23rd September 2-4 pm (UK time)
Rachel Skyner hots a meeting on web apps for fragment-based drug discovery, fragalysis and more.

21st October 2-4 pm (UK time)
KNIME workshop from Daria Goldmann. The KNIME Analytics Platform is the open source software for creating data science. Intuitive, open, and continuously integrating new developments, KNIME makes understanding data and designing data science workflows and reusable components accessible to everyone. KNIME can be downloaded here.

16th November 3-5 pm (UK time)
A PDB workshop to coincide with PDB 50th anniversary. The Protein Data Bank archives information about the 3D shapes of proteins, nucleic acids, and complex assemblies that helps students and researchers understand all aspects of biomedicine and agriculture, from protein synthesis to health and disease.

18th November 3-5 pm (UK time)
Mol* workshop. Mol* is a modern web-based open-source toolkit for visualisation and analysis of large-scale molecular data.

Under Discussion for 2022
Kathy Skopelitou has offered to run a workshop. More details to follow in due course.

Other Suggestions

- Molecular Dynamics
- OpenMM
- NAMD & VMD
- QM
- PSI4
- NWCHEM
- GAMMESS
- Waterswap
- Waterswap
- WaterDock
- Molecular Graphics
- MGLTools
- PyMOL (possibly some more advanced tutorials)
- Docking
- AutoDockTools
- ADFR & ADCP (AutoDock Flexible Residues & AutoDock CrankPep)
- Raccoon2

- FEP
- Retrosynthesis https://rxn.res.ibm.com
- Cheminformatics
- Indigo & Indigo ELN
- https://lifescience.opensource.epam.com/indigo/index.html
- https://lifescience.opensource.epam.com/indigo/eln/index.html

---

Reflections on a rewarding career: Kurt Zielenbach's 37 Years at CAS

Contribution from Kurt Zielenbach, email: kzielenbach@gmail.com

Left: Kurt with the late Bob Massie and original SciFinder Product Development Manager, Kirk Schwall

I am a bit of an anachronism. Unlike many professionals, who periodically change jobs and locations, I live in the same town I grew up in, and for 37 years, I worked for a single employer: a very rewarding career with the CAS division of the American Chemical Society. Within that span of time, I was able to pursue several distinct opportunities, first as a technologist, then as a product developer/manager and finally as a marketing strategist. This somewhat unique set of circumstances provides me with a somewhat unique perspective over a dramatic shift in technology, business practices/models and cultural issues affecting the chemical information industry. My recent retirement has provided me time to reflect on all of this and in turn offer you some of my humble observations.

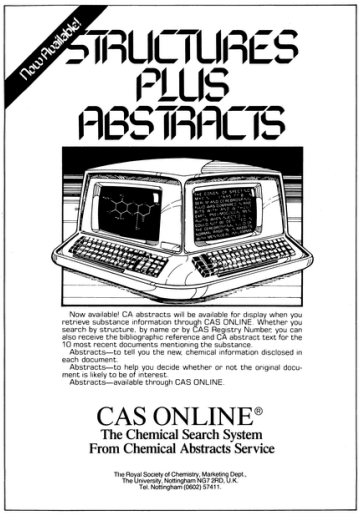
High Tech in the Print Era

I began my time with CAS in 1984 as a "computer programmer". This was an exotic and relatively new field at the time, having only established itself as a career path in the 1970s. Most of my colleagues at the time were math majors, as academic programs in computer science really didn't begin to flourish until about this time.

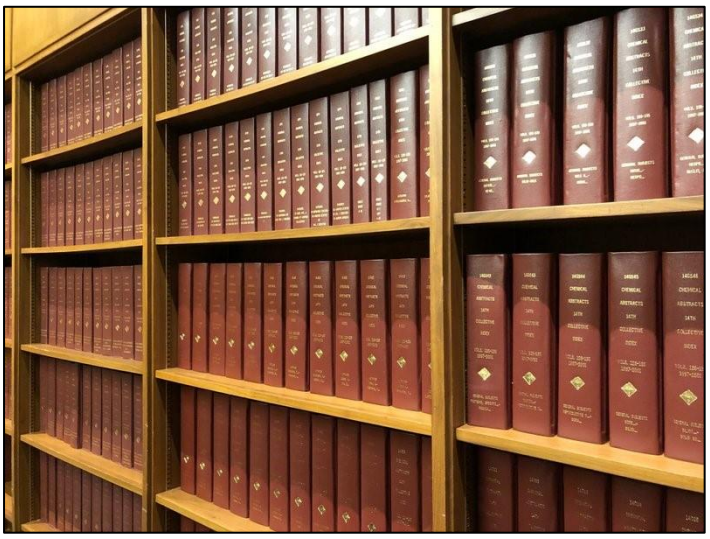
CAS being the technology forward company it was (and continues to be), had begun automating its business extensively in the 1970s, and by the time I arrived, sophisticated systems such as the CAS Registry™ were already in place, largely to support our print products, chiefly CA (Chemical Abstracts), but also various search indices. The printed products filled (literally) the world's scientific libraries both in industry and academia, where scientists would literally spend weeks performing their literature review and due diligence prior to going into the lab, writing a dissertation, or applying for a grant.

The computer programming language de jour was IBM Basic Assembler language which drove the data processing activities of several IBM mainframe computers in the CAS data center. It is amazing to think that the CPU capacity of these behemoth machines pales in comparison with the processing power available on a

mobile device today. Code was written, compiled, and debugged on dedicated terminals and batch jobs that would queue up for execution on the data center mainframes. If you wished to print your code, this was handled by a very large printer, also in the data center. Paper copies of the entire corpus of code were printed on pages of pages of green-bar paper and housed in a large library called Central Documentation for reference, but data processing clerks were also on hand to make periodic deliveries of your printouts throughout the day in grocery-like carts throughout the campus from the big printers in the data center.



CA advert in an RSC journal



A Printed CA collection in all its glory

This set up was probably typical in many ways in any data driven tech-oriented company, but some rather interesting and unique things were also cooking. CAS was already dabbling with pioneering newer operating systems like Unix and revolutionary open packet switching networking technologies like TCP/IP. While I am a bit fuzzy on the details now, I believe there was some limited connection to early networks such as DARPANET or NSFNET. A vivid memory I have was gaining access to this previously unheard-of thing called an email account. It was only usable within the internal CAS network of colleagues, but it quickly became a highly efficient and well used communication tool. Little did I know just how much of a fixture email would become in all our lives.

The Rise of Searchable Online Databases

Another notable new emerging technology at the time at CAS was a pre-internet networked online search system. Customer searchers for the first time could directly search digitised chemical and reference information that previously only existed in print formats in a new service aptly named CAS Online. Users that were able to acquire a very specific device called an HP 2647A terminal could actually graphically select and organise 2D structure features using a tablet and stylus and search the resulting structure against the CAS Registry™. This was truly revolutionary at the time, unleashing incredible new applications and efficiencies for users in how they could consume chemical information that we now take for granted. Initially only available to limited set of customers and only with access to a subset of the CAS Registry, more and more content was added, and the system evolved into what we now know as STN™, a searchable platform of hosted scientific and technical databases from a variety of sources, offered globally through a partnership with FIZ Karlsruhe of Germany and the Japan Science and Technology Agency (JST).

Systems like STN, Dialog and Orbit quickly became the standard means to access chemical information. However, the tools required to effectively and efficiently search them required significant training and repeated usage. Chemists were not able to invest the time and energy into mastering these tools as they wanted to be in the lab making compounds. Academic and corporate science librarians rose to the challenge, with many of them evolving into search specialists where one of their primary responsibilities was performing literature reviews and archival searches on behalf of the research chemists in their organisation. This new intermediary role was a game changer. The time required to survey the literature, prepare for the lab, and

advance research projects was greatly reduced allowing chemists to zero in on the journal articles they needed, and the information professional became a fixture in the library.

My time and attention were soon applied to working with a couple of large chemical companies that hired CAS to host their corporate chemical registries and technical document repositories over a private network, leveraging the database building software as well as the search and retrieval software that served as the engine for STN. As a project lead, I was able to frequently travel to the customer's corporate research headquarters to work with them to understand their information needs and manage and introduce enhancements for further development. This was a monumental step in my professional development. With the ability to connect directly with customers, I had to become much more knowledgeable about our business, our content, and our technologies. I absolutely loved it and it completely changed my career trajectory.

Taking CAS back to the Chemist

While the effective partnership between research scientists in the lab and information professionals in the library was fruitful, it was not without its challenges. Having to send off research questions to a 3rd party information professional would sometimes take the researcher out of their groove, the stops and starts inherent in the workflow making it harder at times to have that *aha* moment that inspires so much innovation. Frequently, search results would arrive leading the researcher to realise that they had asked the wrong question or even more frequently inspiring many additional questions, causing the entire process of sending off information requests and waiting for results to begin again. This could be frustrating of course and just the act of having to send off the questions at all became a barrier at times, with only the most project critical questions being asked. Our research at the time found that for every question being sent out for a search, 6 or 7 questions were not.

What happened next was enabled by a perfect combination of technology, some good ideas and a new management team at CAS that was eager to invest in and launch new products. The '90s brought us the birth of the world wide web and a new generation of increasingly computer literate research scientists with the means to connect with others in a way never before possible. PC technology was taking off with high powered computers available on every desktop and lab and increasingly connected to this new internet "thing". CAS had a new leader in Bob Massie who had brought in a new leadership team that was highly focused on diversifying the CAS product portfolio. Management recognised that this confluence of technologies and information trends opened the door wide open for a new paradigm in cheminformatics, re-establishing a direct relationship between CAS content and the chemists it was born to serve.

Some prototypes emerged from the CAS research department that comprised of a super simple graphical UI search function designed to address a few primary research tasks frequently employed by research scientists. A new product development function had just been stood up at CAS and some intrepid product managers went on the road to show off this prototype and the rest is history. The product idea was pitched to management as technology that would do for research scientists what the newly created software package TurboTax was doing to de-complicate tax preparation.

The product was greenlit, and a development project called Artemis was established with its own software development team, which I was fortunate enough to be a part of. My responsibility was to turn researcher needs into product requirements that could be turned into code by the engineering team. Over the next year and a half, I was able go on the road taking various iterations of the evolving software to chemistry teams pledged to partner with our efforts from several large pharma and chemical companies. This culminated in the launch of SciFinder® in 1995 with a promise to change the way scientists conducted research. It is hard to believe at that time with SciFinder, that scientists were able to ask basic questions and directly review and prioritise answers, all in a span of minutes. Furthermore, even way back in 1995, a scientist could directly download a digital rendering of an ACS journal article to their PC.

This was truly the most treasured portion of my career - learning how research chemists and biochemists tick: what motivated them, what troubled them, and working collaboratively with them to discover ways we could overcome information challenges together through SciFinder.

As we moved into this new century, I took on the role of product manager, helping to manage the SciFinder development roadmap, shepherding it from desktop software to a browser-based solution, and finally helping to evolve it to its latest incarnation, SciFinder^{n}.

Above: an original SciFinder program image

Information Solutions

In more recent years, as CAS under the leadership of its new president Manny Guzman repositioned itself as a full-fledged solutions provider with a rich portfolio of complementary information products and customised services for different phases and scientific and business roles in the innovation pipeline, my particular role also evolved to a solutions marketing strategist. I found it very rewarding to take all the learnings of my SciFinder experiences with researchers and their industries and actualise it through modern digital marketing technologies and techniques to better communicate with research managers and scientists. This allowed them to know CAS could help them solve problems, allowing them to advance their research more quickly and confidently.

Looking back at my 35+ years with CAS, I am truly in awe of my good fortune. I was able to be near the heart of the biggest information innovation since the invention of the printing press with the internet, as well as the mapping of the human genome and all its implications for healthcare research. And I was able to experience these monumental paradigm shifts with an incredibly smart and dedicated cast of colleagues at CAS and ACS, working together to provide essential support to researchers as they tackle and solve the most vexing and important world problems, truly delivering on the ACS mission of improving people's lives through the transforming power of chemistry.

Celebrating the 50th anniversary of the Protein Data Bank

Contribution from Christine Zardecki, RCSB Protein Data Bank, email zardecki@rcsb.rutgers.edu

PROTEIN DATA BANK 50 Years

In the late 1950s, scientists began to decipher the 3D shapes of proteins at the level of individual atoms. The implications of these new structures for science inspired the new field of Structural Biology. Imagining the potential research enabled by archiving and sharing data from these experiments moved the scientific community to action.

In June 1971, a symposium on Structure and Function of Proteins at the three-dimensional level was held at Cold Spring Harbor Laboratory. With its lively conversations, debates, and planning for the future, that meeting defined the beginning of the Protein Data Bank (PDB) as an archive for the experimentally-determined 3D structures of biological macromolecules. Later that year, the Protein Data Bank archive was established as the first open-access, digital resource for biology.

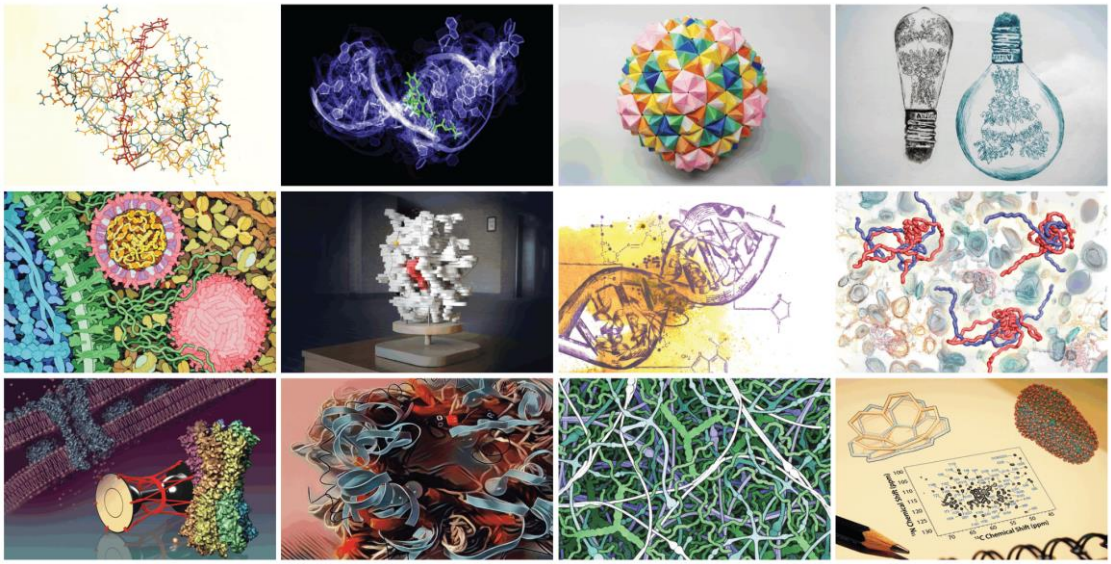
After starting with just 7 crystal structures, the archive today contains and supports online access to ~180,000 biomacromolecular structures determined via macromolecular crystallography, Nuclear Magnetic Resonance spectroscopy, and 3D Electron Microscopy by researchers from around the world. From its inception, the PDB has embraced a culture of open access, leading to its widespread use by the research community. PDB data are used by hundreds of data resources and millions of users exploring fundamental biology, energy, and biomedicine.

In recognition of the importance of long-term preservation of biostructure data, the Worldwide Protein Data Bank (wwPDB) partnership was established in 2003 to manage the PDB archive and ensure that the PDB is freely and publicly available to the global community. It consists of organisations that act as deposition, data processing and distribution centers for PDB data. Current members include the RCSB Protein Data Bank, Protein Data Bank in Europe (PDBe), Protein Data Bank Japan (PDBj), and Biological Magnetic Resonance Data Bank (BMRB).

To commemorate and celebrate 50 years of the PDB, the wwPDB is organising multiple events in 2021 (wwpdb.org/pdb50):

- The inaugural PDB50 event was held virtually in May 2021
- Transactions Symposium 2021: Function Follows Form: Celebrating the 50th Anniversary of the Protein Data Bank (July 30-31, 2021)
  This virtual event is part of the Annual Meeting of the American Crystallographic Association
- Bringing Molecular Structure to Life: 50 Years of the PDB (October 20-22, 2021) Virtual EMBL Conference
- Royal Society of Chemistry PDB Workshop (Nov 16 and 18, virtual)
- Learning from 50 years of the Protein Data Bank: A satellite symposium of the Biophysical Society of Japan (Nov 25-27, 2021)

Visit wwpdb.org/pdb50 for updates and related materials.



Images from the wwPDB's 2021 calendar highlight milestone structures from the PDB archive: lysozyme, HIV-1 TAR RNA, rice dwarf virus, glutamate receptor, Zika virus, cytochrome c, DNA, disordered proteins, gap junction, hemoglobin, collagen, and an HIV capsid.

PDBx/mmCIF News

PDB users and related software developers should be aware of upcoming developments and plans related to the distribution of PDB data. Announcements are made at wwpdb.org.

Modifications to support for SHEET and ligand SITE records in June 2021

In 2014, PDBx/mmCIF became the PDB's archive format and the legacy PDB file format was frozen. In addition to PDBx/mmCIF files for all entries, wwPDB produces PDB format-formatted files for entries that can be represented in this legacy file format (e.g., entries with over 99,999 atoms or with multi-character chain IDs are only available in PDBx/mmCIF).

As the size and complexity of PDB structures increases, additional limitations of the legacy PDB format are becoming apparent and need to be addressed.

Defining complex SHEET records

Restrictions in the SHEET record fields in legacy the PDB file format do not allow for the generation of complex beta sheet topology. Complex beta sheet topologies include instances where beta strands are part of multiple beta sheets and other cases where the definition of the strands within a beta sheet cannot be presented in a linear description. For example, in PDB entry 5wln a large beta barrel structure is created from multiple copies of a single protein; within the beta sheet forming the barrel are instances of a single beta strand making contacts on one side with multiple other strands, even from different chains.

This limitation, however, is not an issue in the PDBx/mmCIF formatted file, where these complex beta sheet topologies can be captured in *_struct_sheet*, *_struct_sheet_order*, *_struct_sheet_range*, and *_struct_sheet_hbond*.

Starting June 8th, 2021, legacy PDB format files will no longer be generated for PDB entries where the SHEET topology cannot be generated. For these structures, wwPDB will continue to provide secondary structure information with helix and sheet information in the PDBx/mmCIF formatted file.

Deprecation of _struct_site (SITE) records

wwPDB regularly reviews the software used during OneDep biocuration. The *_struct_site* and *_struct_site_gen* categories in PDBx/mmCIF (SITE records in the legacy PDB file format) are generated by in-house software and based purely upon distance calculations, and therefore may not reflect biological functional sites.

Starting in June 2021, the in-house legacy software which produces *_struct_site* and *_struct_site_gen* records will be retired, and wwPDB will no longer generate these categories for newly-deposited PDB entries. Existing entries will be unaffected.

Consistent Format for Validation and Coordinate Data

wwPDB validation reports are now provided in PDBx/mmCIF format for all new depositions in OneDep. This change makes validation data more interoperable with the PDB archival format. Data are more logically and better organised in the PDBx/mmCIF reports, and therefore more "database-friendly" than the report in XML format. PDBx/mmCIF-format validation reports for newly released and modified entries will be distributed through the PDB and EMDB Core Archives.

The new PDBx/mmCIF reports are easier to interpret. They contain a high-level summary and offer easier access to residue-level information. Data are provided at multiple levels: entity, chain-specific, and even at the individual residues. For example, it is more straightforward to obtain the total number of clashes. The corresponding validation dictionary is available at mmcif.wwpdb.org/dictionaries/mmcif_pdbx_vrpt.dic/Index. Examples of PDBx/mmCIF validation reports for X-ray, 3DEM, and NMR are publicly available at GitHub.

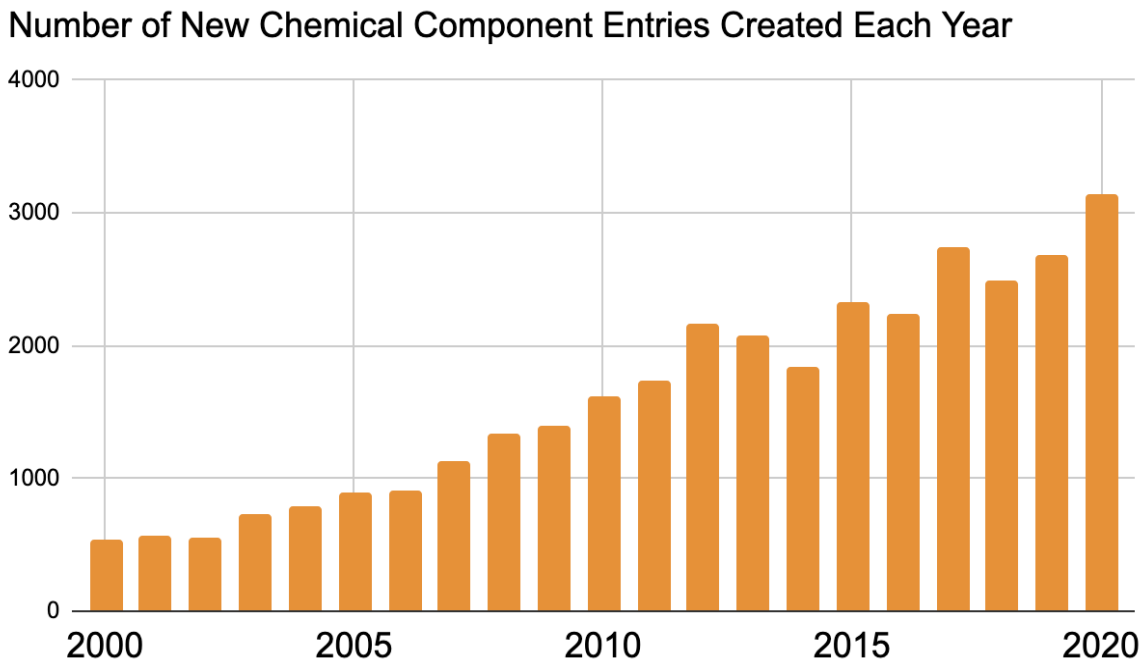
PDBx/mmCIF validation reports will be provided for the full PDB and EMDB archives once archival validation recalculation is performed.

wwPDB strongly recommends all PDB users and software developers adopt this format for future applications.

Future Planning: Entries with extended PDB and CCD ID codes will be distributed in PDBx/mmCIF format only

wwPDB, in collaboration with the PDBx/mmCIF Working Group, has set plans to extend the length of ID codes for PDB and Chemical Component Dictionary (CCD) ID entries in the future. Entries containing these extended IDs will not be supported by the legacy PDB file format.

CCD entries are currently identified by unique three-character alphanumeric codes. At current growth rates, we anticipate running out of available new codes in the next three to four years. At this point, the wwPDB will issue four-character alphanumeric codes for CCD IDs in the OneDep system. Due to constraints of the legacy PDB file format, entries containing these new, four-character ID codes will only be distributed in PDBx/mmCIF format. The wwPDB will begin implementation of extended CCD ID codes in 2022.



In addition, wwPDB also plans to extend PDB ID length to eight characters prefixed by 'PDB', e.g., pdb_00001abc. Each PDB ID has a corresponding Digital Object Identifier (DOI), often required for manuscript submission to journals and described in publications by the structure authors. Both extended PDB IDs and corresponding PDB DOIs, along with existing four-character PDB IDs, will be included in the PDBx/mmCIF formatted files for all new entries by Fall 2021.

For example, PDB entry 1ABC will also have the extended PDB ID (pdb_00001abc) and the corresponding PDB DOI (10.2210/pdb1abc/pdb) listed in the _database_2 PDBx/mmCIF category.

```
loop_
_database_2.database_id
_database_2.database_code
_database_2.pdbx_database_accession
_database_2.pdbx_DOI
PDB              1abc              pdb_00001abc              10.2210/pdb1abc/pdb
WWPDB D_1xxxxxxxxx ? ?
```

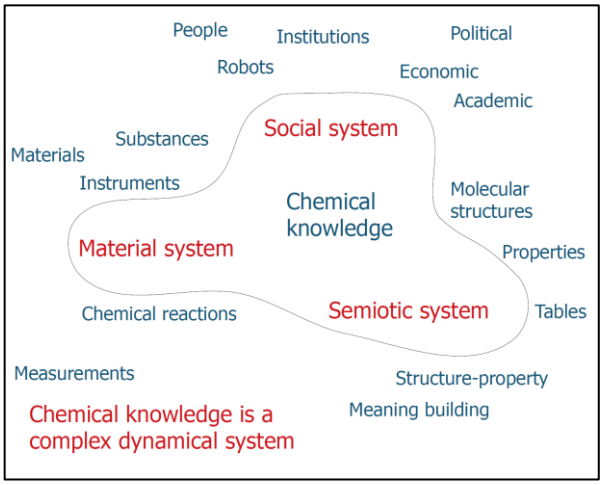
Once four-character PDB IDs are all consumed, newly deposited PDB entries will only be issued extended PDB ID codes, and entries will only be distributed in PDBx/mmCIF format.

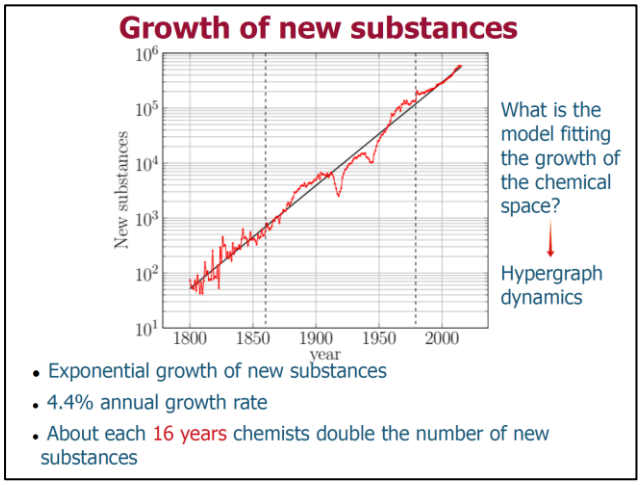
wwPDB is asking PDB users and related software developers to review code and begin to remove such limitations for the future.

Meeting: Computational Approaches to the History of Chemistry

Contribution from Jürgen Jost & Guillermo Restrepo, Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany, email: Guillermo.Restrepo@mis.mpg.de

On 22-23 March 2021 Jürgen Jost and Guillermo Restrepo organised this online meeting, which was hosted by the Max Planck Institute for Mathematics in the Sciences. There were nine forty-minute lectures and two discussion sessions. The aim of the meeting was to discuss the pros and cons of computational approaches to the history of chemistry, the available data, and the required data structures for further computational studies, as well as to present several case studies.





The meeting was motivated by several factors, for example that the increasing amount of data and of computational power are making computational approaches an integral part of the historians' tools. Likewise, that computational history, besides providing new ways to solve historical questions, allows for asking and solving novel questions related to large scale patterns. As noted by the organisers, chemistry, being the science with the largest output of publications associated to its exponential growth of new substances, is therefore not short of data, which is currently well structured in electronic databases and which constitutes an important source of information for historical studies. However, as is true for all sciences, chemistry is more than its material system and it entails other aspects that are not so well historically recorded in electronic form. Discussing the data and data structures of these non-material systems of chemistry was a further motivation for this meeting.

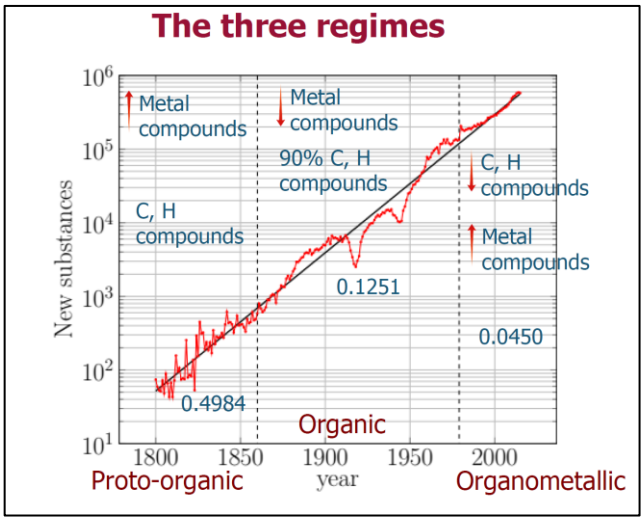
These aims and motivations led us to gather together a group of historians of chemistry, chemists, computer scientists and mathematicians. The speakers were: Gerd Blanke (StructurePendium Technologies GmbH, Germany), Jonathan Goodman (University of Cambridge, UK), Gerd Graßhoff (Humboldt University of Berlin, Germany), Evan Hepler-Smith (Duke University, USA), Jeffrey Johnson (Villanova University, USA), Jürgen Jost (Max Planck Institute for Mathematics in the Sciences, Germany), Carsten Reinhardt (Bielefeld University, Germany), Guillermo Restrepo (Max Planck Institute for Mathematics in the Sciences, Germany) and Peter Stadler (Leipzig University, Germany).

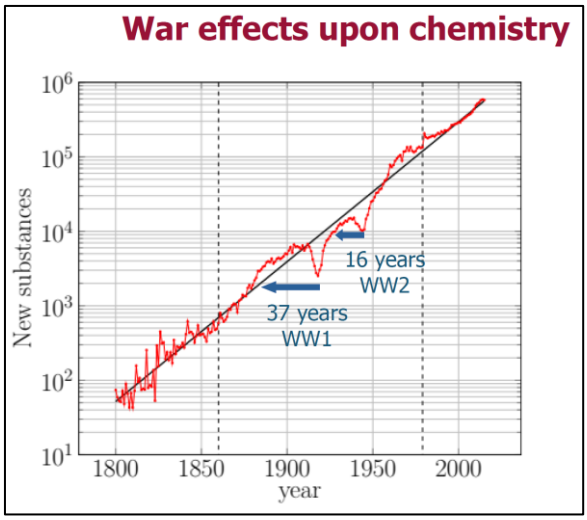
The themes discussed can be framed in the formal setting for the evolution of chemical knowledge introduced by Jürgen Jost and Guillermo Restrepo, where this kind of knowledge arises from the mutual interaction of the social, semiotic and material systems of chemistry. The social system comprises chemists and their institutions and forms of organisation. The semiotic system corresponds to the signs and concepts devised by

chemists to communicate their findings and theorise upon them. Substances, reactions, technologies and apparatus constitute the material system of chemistry. According to Jost and Restrepo, the interaction of these systems can be modelled as a complex and dynamic system, which when unfolding over the history, constitutes an important tool for the history of chemistry.

Carsten Reinhardt provided an example of the interaction of the three systems of chemical knowledge when discussing the rise of chemical instrumentation in the 1950s and its social, semiotic and material conditions. This included new forms of scientific organisation after WW2, novel electronic-based technologies, as well as semiotic objects, such as spectra and software. Evan Hepler-Smith showed how nomenclature systems and methods to store chemical information resulted from the interaction of social interests with the amount and diversity of chemical substances and the different ways of highlighting relevant aspects of these substances. Gerd Graßhoff, when discussing his computational model for the discovery of the urea cycle, emphasised the importance of including semiotic, material and social aspects in those models, represented as the ways of encoding chemical knowledge, the number of substances and reactions and the scientific influences constituting the discovery context.

By analysing some aspects of the social, semiotic and material systems of chemistry in the 19th and 20th centuries, Jeffrey Johnson considered the growth of chemistry institutes and of chemistry staffs, as well as the number of scientific publications and the growth of inorganic and organic chemistry. He found that after WW1 chemistry reduced its growth and he highlighted possible reasons tied to the public perception of chemistry and to the social establishment of chemists. Johnson emphasised the importance of data sources such as the *Chemisches Zentralblatt* for historical studies of chemistry. By the time Johnson conducted his studies this source was in printed form and today is fully available in electronic form, which Johnson takes as an opportunity for computational studies in the history of chemistry. Besides the Chemisches Zentralblatt as a source of information for the material system of chemistry, Johnson discussed the H. Plessner-C von Ferber non digital database for comparative studies of the evolution of the chemical community, which represents a target source to be digitised.





When discussing data for computational studies on the history of chemistry, Peter Stadler emphasised the highly structured data for the material system of chemistry, today available in large repositories of chemical information, which gathers together information on substances and reactions. Stadler showed how a formal setting for chemical reactions, such as regarding them as graph rewriting rules acting upon molecular structures, may enable us to trace the historical rise of new reaction classes and even to solve questions of the sort "what if" by perturbing the data. By this he means, for instance, taking the known substances of an arbitrary past time and applying to them a given set of rewriting rules. This allows for determining what would have been the possible set of substances by the action of the applied reactions.

An instance of the highly structured data of the material system, discussed by Stadler, is the Reaxys© database, which results from the merging of the Gmelin and Beilstein Handbooks and incorporates chemical information from the most salient contemporary chemical literature. The exploration of Reaxys© to pose and solve historical questions was the subject of Guillermo Restrepo's talk, who showed that the number of new

substances taking part in chemical reactions (chemical space) has historically expanded in an exponential way following three clearly distinguishable regimes, with transitions occurring in 1860 and 1980. Likewise, it was discussed how this space has been concentrating, over the history, on substances made of carbon, nitrogen, hydrogen and oxygen. Restrepo also discussed the fixed substrate approach, which has been the disciplinary way chemists have used to combine substances in a rather conservative manner to expand the chemical space.

Regarding the binary relationships between the constitutive systems of chemical knowledge, Hepler-Smith provided an example of the social-semiotic relationship when illustrating how chemical nomenclature has been related to social interests of national chemical communities, for instance at the Geneva congress of 1892. Restrepo provided an example of the social-material interaction through the effect of the WWs upon chemical production. Gerd Blanke illustrated the semiotic-material relationship, where semiotic factors such as the advent of machine learning algorithms, coupled with material factors such as the growing number of chemical reactions, have led to developing further semiotic objects, for instance machine readable representations of chemical reactions. Blanke and Stadler discussed the history of molecular representations and Blanke emphasised their use for encoding information on chemical reactions for chemoinformatic and history of chemistry purposes.

Further insight on the different ways of annotating substances, initially regarded as macroscopic entities undergoing chemical transformations, as noted by Hepler-Smith, and later-on, concentrated on the molecular structure, were the subjects of Blanke, Stadler and Jonathan Goodman's talks. In particular, Goodman presented the advantages and disadvantages of several of these annotation methods and highlighted the semiotic load of molecular representations, as they affect the way chemists regard molecules. He stressed the advantages of RInChI for computational studies related to chemical reactions and mentioned the discussions in the chemical community about what to include and what not to include as part of this annotation procedure for reactions.

Blanke analysed the problems and challenges of annotating chemical reactions, which arise from the different formats and the different emphasis on some particular reaction information of some of those methods. This variety of formats led to the discussion of the suitable information for historical studies and the kinds of formats pertaining to the systems of chemical knowledge. This was the subject of the second discussion during the meeting, where the rich and structured information of the material system of chemistry was highlighted, which contrasts with the isolated databases for the social system, as those presented by Johnson, and with the lack of semiotic databases.

The first discussion was about the pros and cons of computational approaches to the history of chemistry. As pros, the importance of these approaches for detecting large-scale historical patterns, as those discussed by Restrepo, was mentioned. A further advantage was the possibility of reconsidering accepted ideas in the history of chemistry with support on large corpora of data. In this case, the role of synthesis in the expansion of the chemical space can be mentioned, as discussed by Restrepo. It is traditionally accepted that synthesis began after Wöhler's synthesis of urea in 1828, but the evolution of the chemical space shows that synthesis played a major role in the expansion of the chemical space throughout the entire 19th century. A second example challenging accepted ideas involved the historical study of the interplay between the chemical space and the periodic system. Restrepo showed that the size and diversity of the chemical space in the 1840s already provided the salient features of the periodic system of the 1860s, which contrast with the accepted account that the ripe moment to formulate the system was in the 1860s. A further advantage of computational methods involved the possibility of asking new historical questions and even of perturbing the data and observing the resulting temporal effects, as discussed by Stadler.

Some of the disadvantages of the computational approaches discussed involved the lack of stable historical records, as highlighted by Blanke, who made the point that electronic databases, especially if updated regularly, may change the historical record by introducing new data of the past or by correcting annotation errors. Therefore, it was discussed that stable dumps of these databases need to be stored, which requires negotiations with database providers and also a computational infrastructure to store and share this information for the aims of historians of chemistry.

The meeting concluded with the general feeling that computational methods constitute a suitable tool to complement the methods of history of chemistry and that the further advance of these approaches requires interdisciplinary work among historians, chemists, mathematicians, computer scientists and other specialists.

Meeting: Being #CompChemURG – How Diversity Enriches us

Contribution from Hannah Bruce Macdonald, email hannah.brucemacdonald@choderalab.org

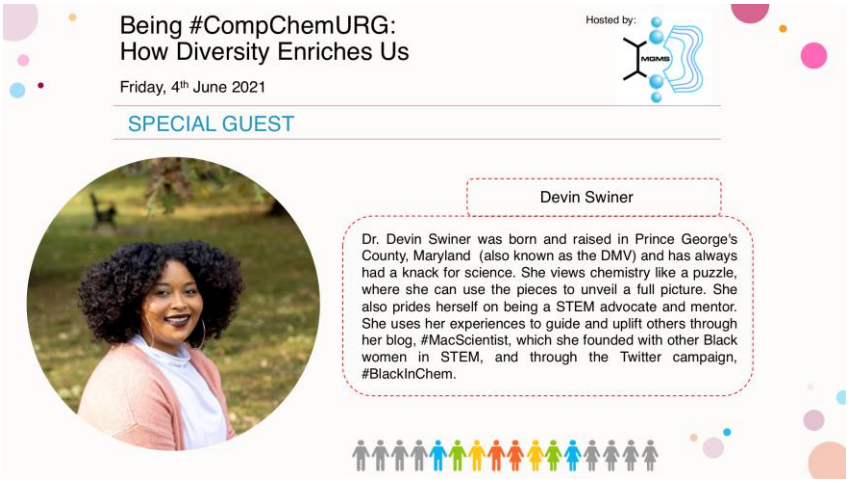
THE BINDING SITE

Resources for underrepresented groups and allies in Computational Chemistry

The inaugural CompChemURG meeting was held as a virtual event on 4th June 2021, with funding and support from the RSC's Inclusion and Diversity Fund and the Molecular Graphics and Modelling Society (MGMS). The scope of the meeting was to discuss the range of diversity there is in our computational chemistry community and initiate an online community (www.bindingsites.co.uk) of underrepresented groups in our field and provide resources.

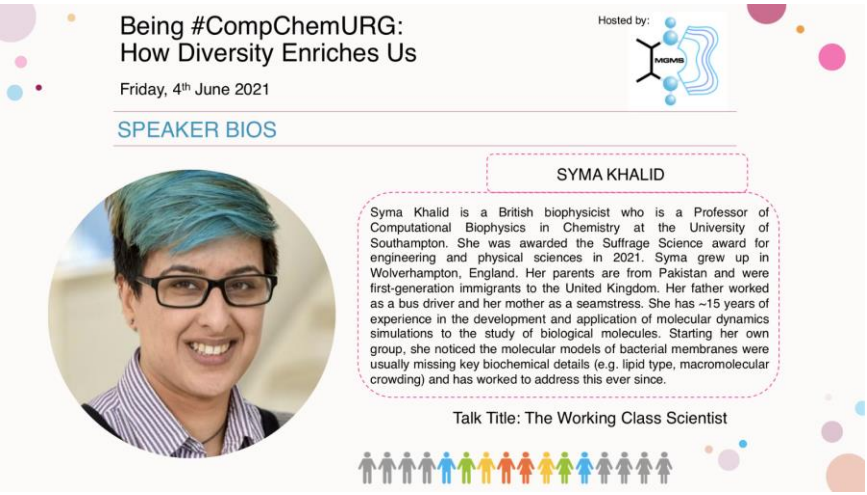
The event had 277 registrants, and was hosted on zoom by MGMS, with technical support from Ewa Chudyk. Holding the conference virtually was not only done because of COVID-19 restrictions, but to open the event to a wider audience and provide the potential of anonymity for attendees. Due to the emphasis on the privacy of attendees, no details of attendees were recorded. There was good online engagement of the event, via the twitter account @CompChemURG and the hashtags #CompChemURG and #CompChemEveryone.

Main session



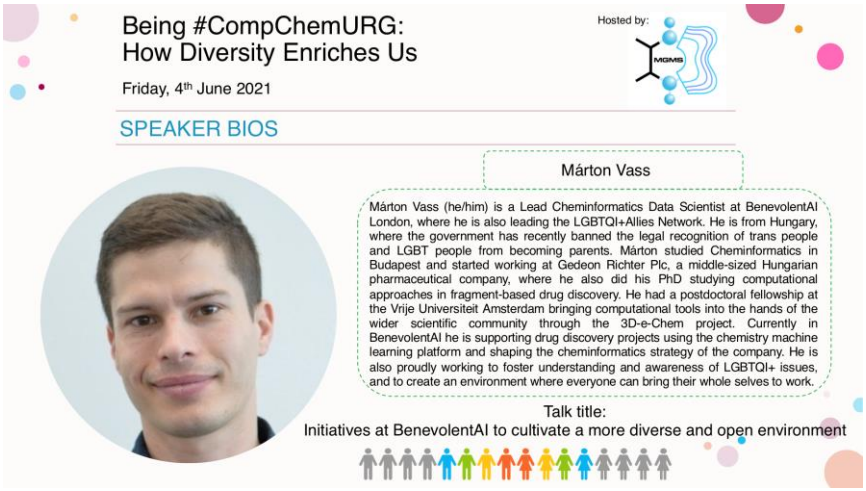
The conference and it's aims and intentions were introduced by the event organisers, Michelle A. Sahai (Michelle.Sahai@roehampton.ac.uk) and Hannah Bruce Macdonald, with an outline of the upcoming talks. The second half of the introduction was given by guest speaker Devin Swiner (@devin_eleven), Co-Founder of #BlackInChem and The Chemist of @MacScientists, who gave an outline of the start of the online #BlackInChem community and the successes of the organisation. The main session of the conference comprised an afternoon of talks, from eight prominent computational chemists in the field as invited speakers, to discuss their broad experiences.

Syma Khalid, University of Southampton: The Working-Class Scientist



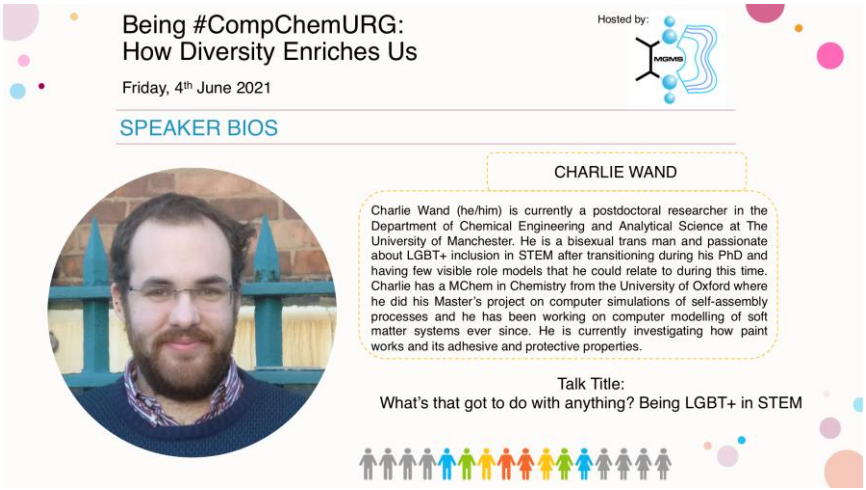
Syma began the talk session, walking us through the different barriers that can be faced by the working-class community, starting from school age, passing through university and beyond. Syma not only provided her own perspective, but an overview canvassed from many of her peers. She both highlighted disparities, but also highlighted indications of positive change in the statistics of backgrounds of those currently entering higher education.

Márton Vass, BenevolentAI: Initiatives at BenevolentAI to cultivate a more diverse and open environment



Márton discussed the efforts of the LGBTQI+Allies network within the industrial setting of BenevolentAI. He described the support the company provided to the group, and the efforts that the group was initiating too. He discussed how allies to the LGBTQI+ community were included, while also creating the safe space for those who need it.

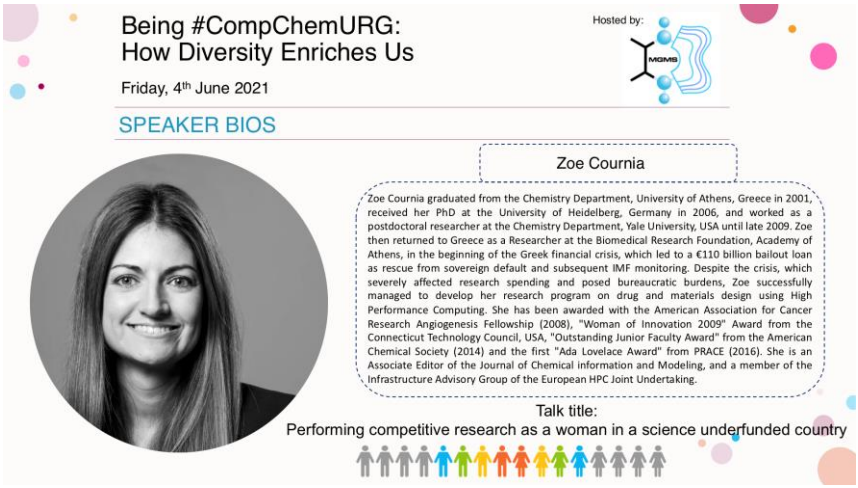
Charlie Wand, University of Manchester: What's that got to do with anything? Being LGBT+ in STEM



Charlie discussed his background, and the difficulties involved in both transitioning as a science undergraduate without role models that looked like him, and his experience with coping with being transgender in a foreign country. Charlie provided a very personal outlook on the difficulties that he has faced and how his life is something that does come to work with him.

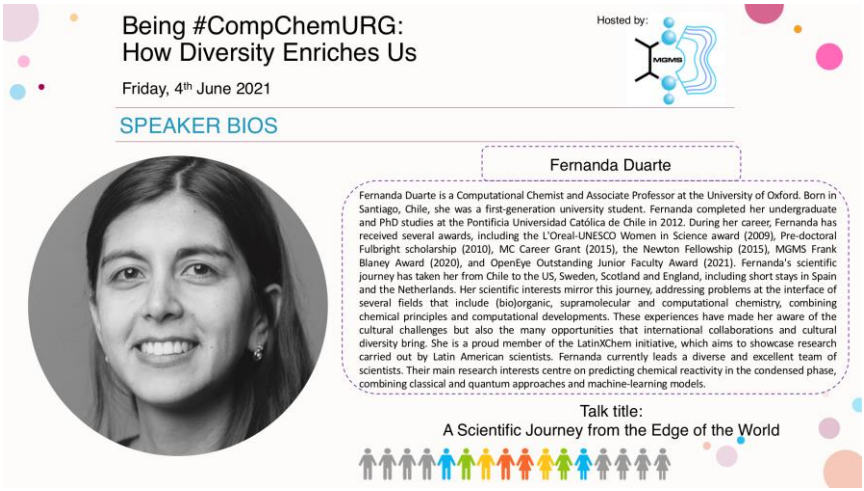
Zoe Cournia, Academy of Athens: Performing research as a woman in a science underfunded country

Zoe discussed the difficulties she faced moving back to Greece to establish her research group during the global financial recession in a hard-hit country. She discussed how at the beginning she acted as her own post-doc and has benefitted from undergraduate and summer research students to establish her lab. Zoe provided details of many non-geographically specific funding sources she has applied for, which are now listed under the 'Resources' section of our website.



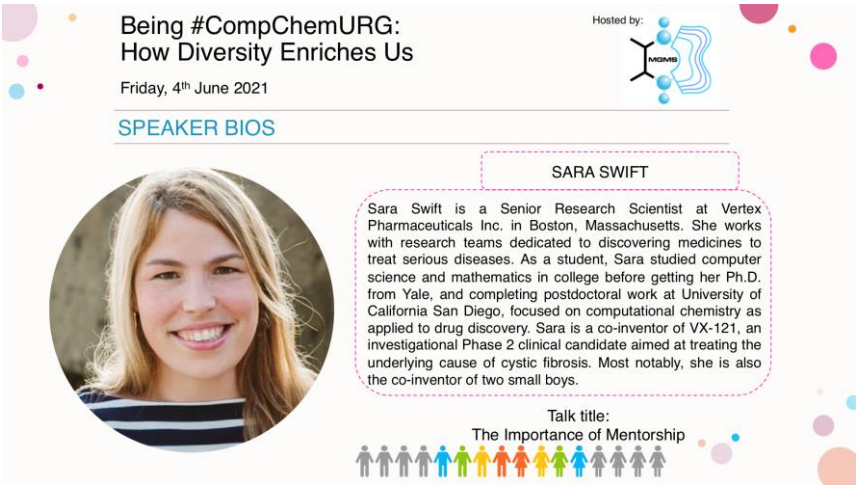
Fernanda Duarte, University of Oxford: A scientific journey from the edge of the world

Fernanda discussed her initial introduction to computational chemistry through discussing publications with mentors, which then began her career moving between many countries. Fernanda discussed how mobility could be a challenge when English is not your first language but described the support she had from supervisors throughout the years, and how global this impact was on her research.

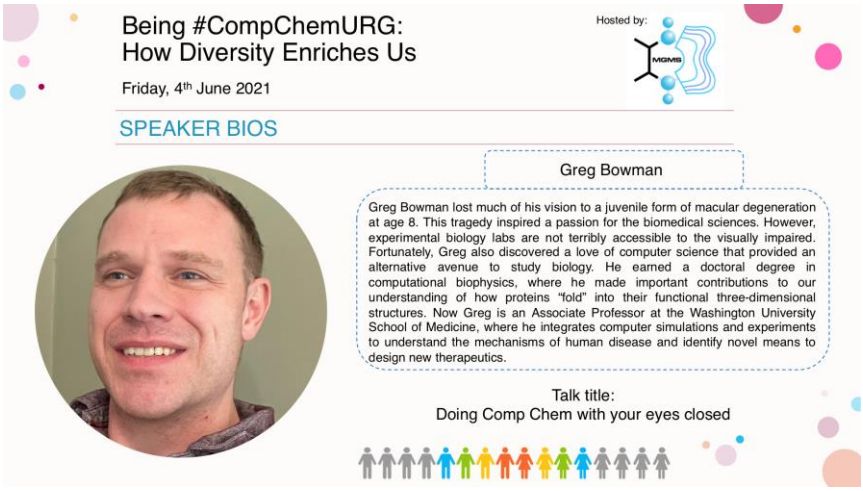


Sara Swift, Vertex: The importance of mentorship

Sara discussed the many roles each of us play in our lives, beyond those we present just at work. She discussed the important roles that mentors have played for her, throughout her career – both her peers and those senior to her. She also discussed the current efforts of equity, diversity and inclusion within the industrial setting of Vertex.

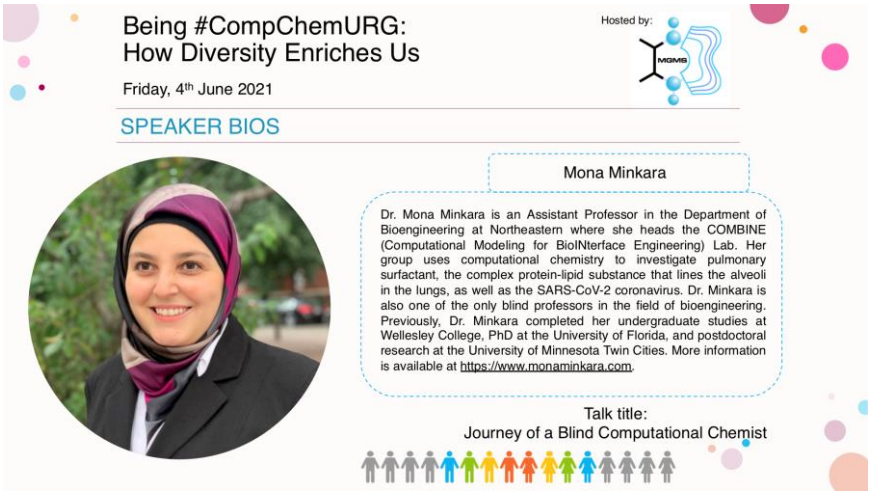


Greg Bowman, Washington University – St Louis: Doing comp chem with your eyes closed



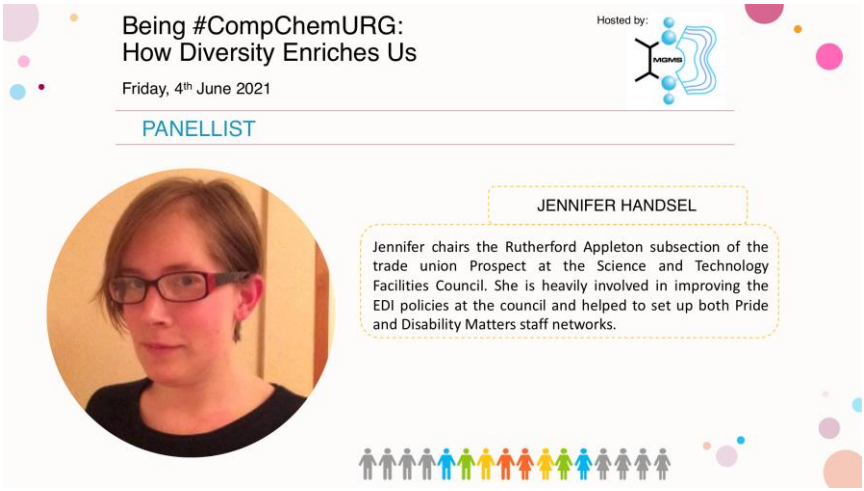
Greg provided the background of his childhood growing up and learning of his visibility issues at a young age which was identified when he was not being able to see the board. Greg showcased the equipment that has helped him to see throughout his life, and he discussed how previous supervisors had been supportive of providing the right equipment for him, e.g. in the form of an extra large monitor. Greg also addressed the difficulty of networking with poor eyesight, and the barriers that name badges and holding eye contact can prove at conferences. He also discussed how painful the paradox can be between your successes and failures and how having a disability (physical or hidden) can affect one's self-esteem and ego.

Mona Minkara, Northeastern University: Journey of a blind computational chemist



Mona provided the final talk of the session, discussing her journey in computational chemistry, and how her vision led to creativity in developing additional methods for analysing trajectory dynamics. Mona also highlighted the importance of supportive mentors within her career.

Panel discussion



The conference was concluded with a panel discussion, where Jennifer Handsel, an EDI expert from the UK Research and Innovation (UKRI) council joined our speakers to discuss questions on the theme "*Finding your community in computational chemistry*". The panel discussion lasted for an hour and addressed questions that were received via twitter in advance of the conference and in the zoom Q+A during the event. After the event, many of the speakers made themselves available to attendees via personal online meeting platforms, to be able to discuss one-on-one with various conference attendees. We are grateful to the speakers, as this provided connection opportunities that can be lacking for virtual conferences.

The popularity of the event, and the feedback received afterwards reflected that the group was addressing an unmet need in the community. The following steps for the group are to develop the 'community' section on

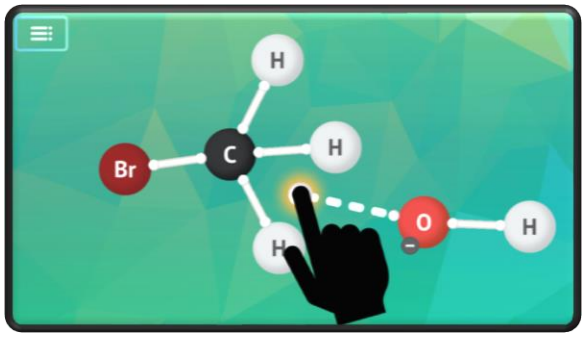
the website, highlighting computational chemists from a diversity of underrepresented minorities, who can join the community via the online form. Work is in progress to establish a mentorship scheme within the community, to forge connections and lessen the barriers of entry to the field. The event marks the beginning of the online community, both on the website (www.bindingsites.co.uk) and the twitter @CompChemURG. If you have any comments or questions, please contact us via either of these mediums, or through emailing compchemURG@gmail.com. The organisers are grateful to the MGMS for their support, and encouraging feedback, and to the RSC Inclusion and Diversity fund for supporting the event.

Visualising Organic Chemistry Mechanisms in 3D with Alchemie

Contribution from Gianna Manchester, email: gianna@alchem.ie

Since its founding in 2016, Alchemie's company mission has been to help students gain an instinctual understanding of electron movement in organic chemistry mechanisms.

The first step toward accomplishing this goal was creating the Mechanisms app. Designed primarily for touch-screen devices, the app emphasises mechanistic reasoning and the flow of electrons by allowing students to use their fingers to move electrons to make and break bonds and see the resultant 2-dimensional structures form. Through a Small Business Innovation Research grant from the U.S. National Science Foundation, we



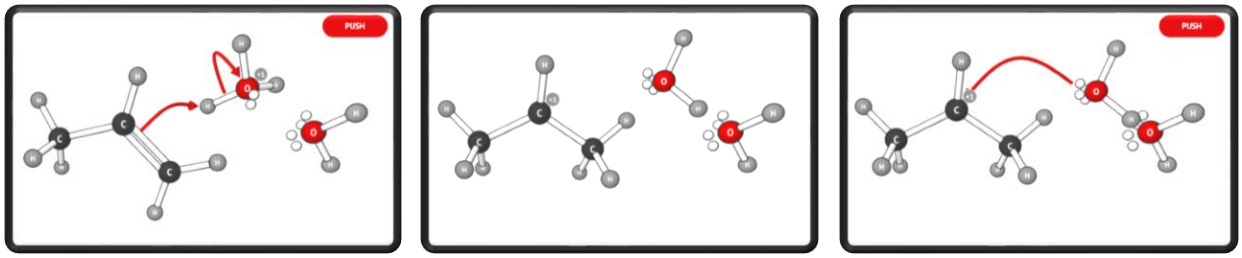
Left: The Mechanisms app permits users to drag electrons with their finger to make and break bonds

designed 265 mechanism puzzles covering a full year of organic chemistry, all accompanied by hints, instant feedback, and videos to guide students' learning. Used in over a hundred schools worldwide, the Mechanisms app has allowed Alchemie to aid students in developing their mechanistic reasoning skills and understanding of electron movement. In pursuit of further striving towards our company mission, we have since developed a new mechanism tool, but also because there are two key components the Mechanisms app lacks: the ability to show stereochemistry and the 3-dimensional nature of molecules and, ironically enough, arrows.

Although a bit controversial to some, the Mechanisms app does not include arrows for two reasons: (1) Arrows are an artifact of drawing mechanisms on paper as a way to represent the movement of electrons. In the app, users can perform that movement, thus eliminating the need for such representation. (2) Having students move the electrons themselves opposed to drawing arrows removes the disconnect in understanding the representation. Anecdotal evidence from students and instructors shows that students understand the movement of electrons better when using the Mechanisms app. This was our steadfast philosophy for years; that is until we began developing our latest tool, Mechanisms 3D.

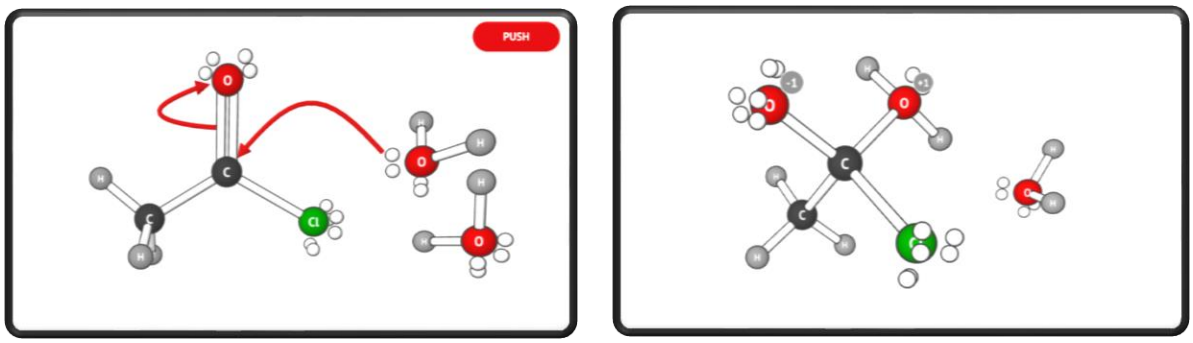
The plan for the next generation mechanisms tool is to display the molecules as three-dimensional structures. Unlike the Mechanisms app, Mechanisms 3D would allow consideration of stereochemistry and proper alignment of molecules during a reaction. Mechanisms 3D was built using our existing tool, Modeler, as a backbone. We partnered with WebMO to integrate molecular modeling into Modeler and Mechanisms 3D to accurately show the geometry and layout of molecules. To make Mechanisms 3D, we added in the ability to move electrons, animate products forming, and record the sequence in which electrons were moved.

When first designing Mechanisms 3D, it was not our intention to include arrows. We still believed that arrows are an artifact of drawing mechanisms on paper and within the digital interactive, users can move the electrons themselves, thus eliminating the need for arrows. But in an early prototype, a cursor was added to draw attention to the electrons moving in a replay of a mechanism. Ironically, the content team was amazed by what seemingly resembled an arrow. After much debate and iterative design, the Alchemie team collectively decided on including arrows in Mechanisms 3D with hopes that it would help students make the connection between the arrow and the movement of electrons that it represents. It is now our mission to not only help students understand the movement of electrons, but also to better connect the dynamic electron movement implied by curved arrows in static drawings.



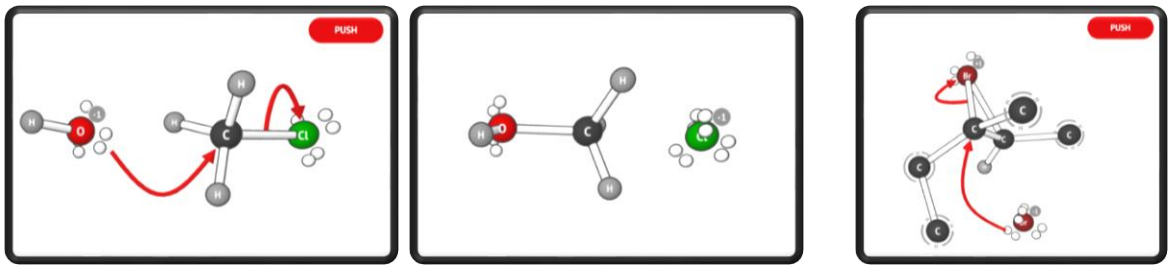
Above: Mechanisms 3D allows users to draw arrows then click PUSH to see an animation of the electrons moving to form the product

In Mechanisms 3D, arrows are drawn by selecting electrons, either a bond or lone pair, and dragging the arrowhead to the atom or bond to which the electrons will go. Once the arrow or arrows for that step are drawn, the "PUSH" button is clicked and an animation of the electrons moving ensues thus forming the resultant product. The animation of electrons moving allows students to visualise the movement and understand how the breaking and making of bonds results in the formation of the product. This directly demonstrates what the arrow represents and emphasises electron movement which should be the focus of mechanisms. Oftentimes when drawing mechanisms on paper, students become so focused with redrawing structures that they lose sight of what happened to get from one structure to another.



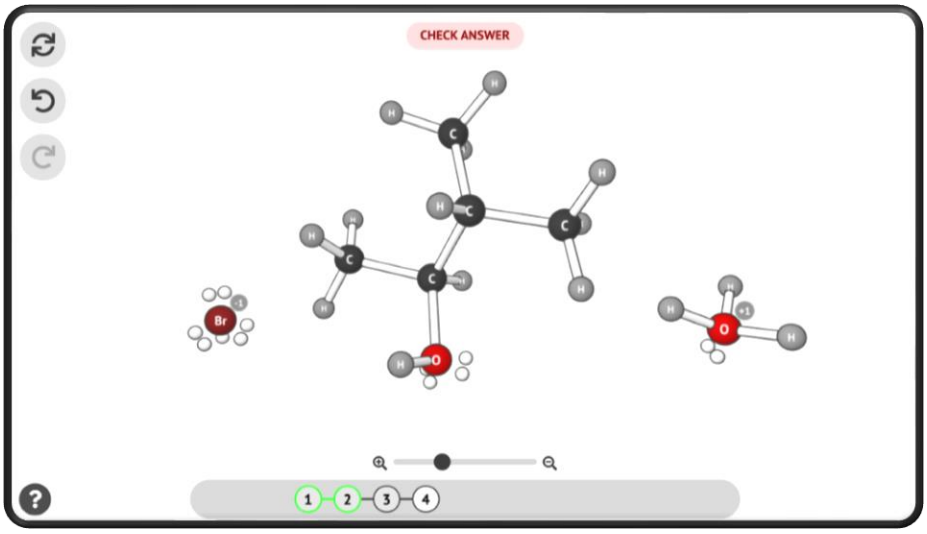
Mechanisms 3D shows the change in spatial arrangement of atoms as the reaction progresses. The 3D nature is particularly helpful when looking at the change that occurs in the formation of the tetrahedral intermediate of a nucleophilic attack on a carbonyl.

Since Mechanisms 3D shows the 3-dimensional structures of molecules, students are able to see the stereochemistry of the reaction. With nucleophilic attacks on carbonyls, for example, the formation of the tetrahedral intermediate becomes quite apparent. Likewise, in SN2 reactions, students can position the nucleophile for the backside attack and watch the inversion of the stereocenter happen as the concerted mechanism ensues. Or even see why bromine adds anti in the bromination of an alkene opposed to simply memorising how to draw the wedges and dashes in the product.



Mechanisms 3D allows users to see the arrangement of molecules in 3D, demonstrating the importance of positioning in backside attack of SN2 mechanisms and anti addition in bromination of alkenes.

A timeline of steps at the bottom permits users to navigate through the mechanism, revisit previous steps, and see a replay of the mechanism. In the replay, the arrows briefly appear before the animation of the electron movement and product formation. Unlike on paper, the digital nature of Mechanisms 3D allows students to experiment with ease by trying out moves, quickly seeing the results, and going back if necessary. Similarly to using paper, students are allowed to make mistakes such as forming five bonds to carbon. However, because the result of the arrows drawn is automatically formed, students cannot skip steps within the mechanism to miraculously get to certain intermediates or products. Rather, every arrow must be drawn in order to get from one step to another, thus reinforcing the need for mechanistic reasoning and an understanding of how electrons move.



In Mechanisms 3D, students can check their answer and see which steps were correct. For example, the student can see the first step from State 1 to State 2 was correct but moves after that were incorrect. They can return to State 2 and realise they forgot to perform a hydride shift.

Mechanisms 3D was designed as a learning tool to help students visualise and understand electron movement as represented by arrows in mechanisms. Alchemie's comprehensive suite of 3-dimensional interactives all have LTI compatibility ensuring seamless implementation in a variety of ways including digital textbooks, as a demonstration tool, and in Alchemie's new online learning system. All of our learning tools, including Mechanisms 3D, have assessment capability which has potential to revolutionise the way students demonstrate their understanding of mechanisms and structural concepts of chemistry. Happy recording!

Chemical Data Recovery 1: Embedded Chemical Data

Contribution from Kevin Theisen, President, iChemLabs, email: kevin@ichemlabs.com

This follows on from Kevin's contribution to the Winter 2020-21 Newsletter on parallel processing for molecular modelling and is the first in a series of three articles submitted from Kevin Thiesen on the subject of chemical data recovery:

1. Embedded Chemical Data Recovery
2. Chemical Image Recovery
3. Legacy Chemical Data Recovery

We launched ChemDoodle 2D v11.2 on December 4, 2020. Included were powerful new tools for recovering chemical data from files, including Microsoft Office files, which this article discusses in detail.

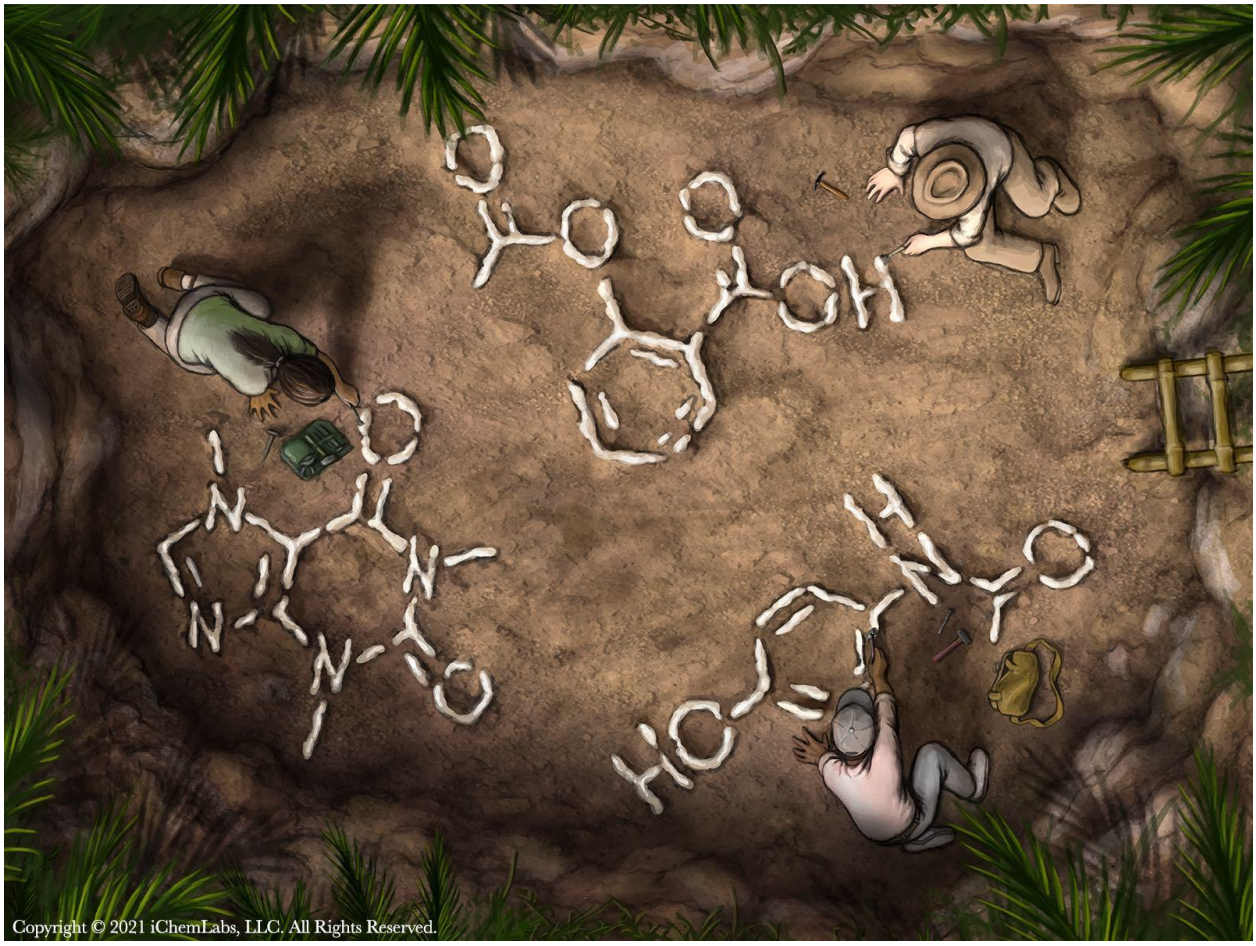


Figure 1: In this artistic rendering, several archaeologists unearth fossils of chemical structures. This is a metaphor for new chemical data recovery features in ChemDoodle 2D.

The handling of chemical data is extremely important, as without it, chemists would not be able to store or communicate chemistry information through computers. Over the past century, many file formats have been proposed to represent chemical information. Several formats have become essential and are implemented by the vast majority of chemical software. These formats include MDL connection table formats (*MOL*, *SDF*, etc.), Chemical Markup Language (*CML*), and *CDX*/*CDXML* formats. There is a fair amount of chemistry software that "read" chemical files. However, most software that read chemical files understand little more than basic molecules and perhaps reactions, while there are a lot of objects and properties defined by the various formats for creating complex chemical data and figures. The goal of chemical file reading in ChemDoodle is to fully understand all objects and properties and to recover the graphics of the chemical figures drawn in other programs in a pixel-perfect manner.

Also prevalent in the professional chemistry community is an application process referred to as "round-trip editing" (fully supported by ChemDoodle and discussed in detail in the ROUND TRIP EDITING section of the ChemDoodle 2D user guide), which allows a user to create data in one program (such as ChemDoodle)

and store the data in another program (such as Microsoft Office). This allows a chemist to easily manage chemical data, which is not an intrinsic data type in other programs like text or images are, without having to save and manage multiple files from different applications. For instance, if a chemist includes a dozen reaction schemes in a word processor, then those reaction schemes can later be accessed and edited through the word processor's application and file.

Professional chemists have come to rely on the convenience of round-trip editing, but it is fraught with risk. On several occasions in the past decade or so, Microsoft Office has dropped round-trip editing support on macOS, leading to many macOS users losing access to their chemical data. Therefore, as a practical consideration when working with niche data like chemical data, your computer and software will eventually change and there is a chance you will be locked out of accessing your data. Some software may become obsolete, the developer may change it, you may switch to a different operating system where software is no longer compatible, or you may no longer be able to afford the products you used in the past. In all of these cases, you will no longer have access to the original chemical data you created. Years of work and effort may be lost in this way.

ChemDoodle includes very thorough tools for recovering chemical data, whether through files, via interactions with other software, or through round-trip editing. In particular, ChemDoodle expertly handles chemical data and can recover embedded chemical data in other applications, regardless of the operating system you are using. This allows you to maintain control and access to your work. This process is very complex, and so we describe the different ways of recovering chemical data as levels. There are 4 different levels of chemical data recovery. ChemDoodle fully supports all 4 levels on Windows, macOS and Linux.

0. Reading chemical files
1. Pasting directly from other chemical applications
2. Pasting embedded data from other (non-chemical) applications
3. Recovering data from Microsoft Office files

Level 0: Reading chemical files

The most basic function for recovering chemical data is to read chemical files. ChemDoodle understands over 30 widely used chemical file formats. As stated in the introduction to this section, the goal of chemical file interpretation in ChemDoodle is to fully and completely handle file contents and to recover the graphics in a pixel perfect manner.

Data and graphics support in ChemDoodle is an immense undertaking. Yet, the ChemDoodle application does this expertly. If you are not satisfied with the way ChemDoodle opens your chemical file, simply send it to us and tell us what you are not satisfied with, and our goal will be to improve it.

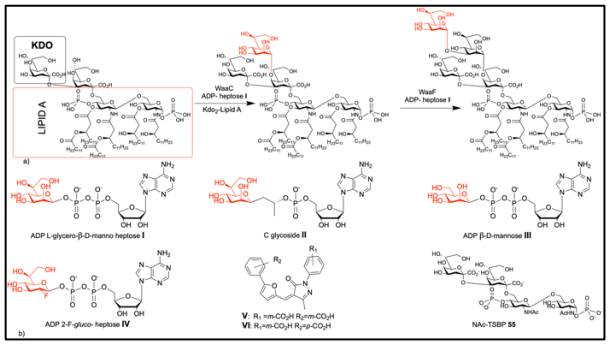


Figure 2a: A complex ChemDraw figure, exported as an image from ChemDraw

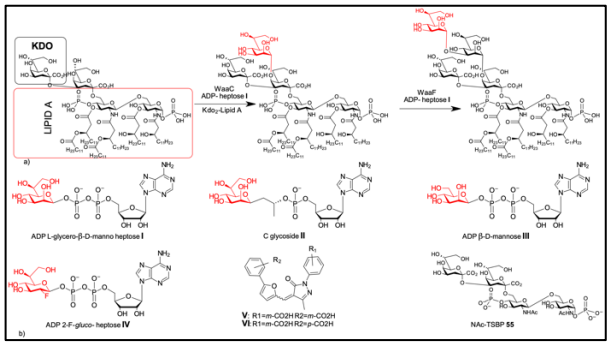


Figure 2b: The same figure, imported into ChemDoodle as a CDX file, and exported as an image from ChemDoodle

Level 1: Pasting directly from other chemical applications

While it may not seem sensible to need two separate chemistry applications or ever to need two different chemical drawing applications, the scope of each application is different, and there are functions ChemDoodle

can perform that other chemical software cannot. For instance, you may want to use ChemDoodle to style bond strokes or generate 3D coordinates. You can copy from other chemistry applications and paste into ChemDoodle on both Windows and macOS. ChemDoodle is compatible with *ChemDoodle Collages*, *ChemDoodle JSON*, *MDLCT* and *MDLSK*, and *ChemDraw Interchange Format* clipboard data on Windows, and is compatible with *ChemDoodle Collages*, *ChemDoodle JSON* and *ChemDraw Interchange Format* clipboard data on macOS.

Level 2: Pasting embedded data from other (non-chemical) applications

Some applications embed chemical metadata into figures pasted into 3rd party applications, in a process referred to as "round-trip editing". The technical procedure for recovering this data is different on Windows and macOS. However, on both operating systems, these figures can be copied from the 3rd party application and pasted into ChemDoodle to recover the original data for further editing without using or having access to the original chemical software that created the figure.

This process will be limited to the operating system the figure was created on. This is because Windows and macOS handle system metadata differently and most programs working on both operating systems (like Microsoft Office) do not convert the data for the other operating system.

As an example, if you have pasted a chemical figure into Microsoft Word on Windows, you will later be able to copy that chemical figure from Microsoft Word on Windows and paste into ChemDoodle for further editing. And if you have pasted a chemical figure into Microsoft Word on macOS, you will later be able to copy that chemical figure from Microsoft Word on macOS and paste into ChemDoodle for further editing. This works for any application correctly storing the chemical metadata, including WordPad on Windows and Pages and Mail on macOS.

If you need to get around this operating system limitation for Microsoft Office, look at the next section, Level 3.

Level 3: Recovering data from Microsoft Office files

ChemDoodle provides a unique solution for searching and extracting chemical figures out of, specifically, Microsoft Office files, regardless of what application or operating system the embedding occurred with. Microsoft Word will be the most common filetype used to store embedded chemical data. This is especially useful if you are a macOS user collaborating with a chemist on Windows (and *vice versa*), or if you used to be a Windows user but switched to Linux, or if round-trip editing is broken by any part of the chain. To recover chemical data using this method in ChemDoodle, you do not need access to other chemistry software, Microsoft Office, or the original operating system the data was created on.

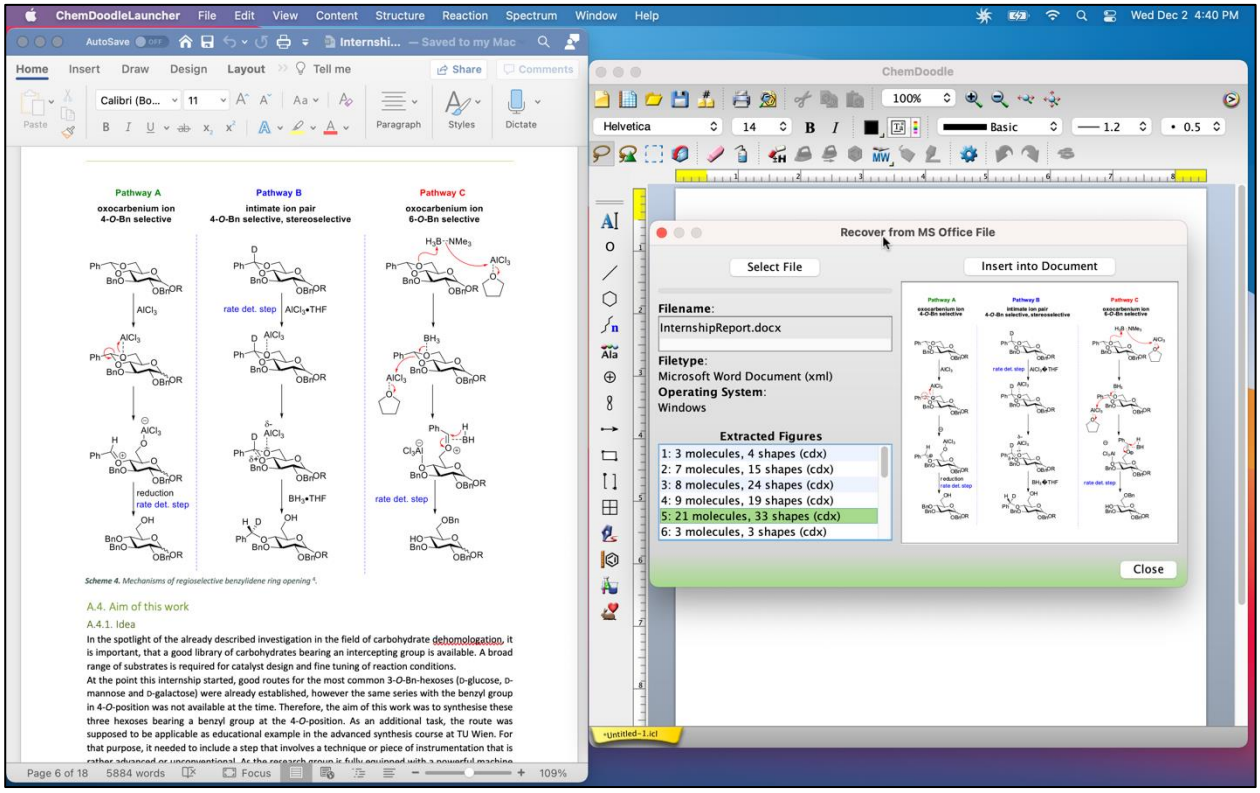


Figure 3: In this image, ChemDoodle on macOS recovers complex chemical figures drawn and pasted from ChemDraw into Microsoft Word on Windows.

Currently, the *XML* files from Word, Powerpoint and Excel can be handled, and ChemDoodle, ChemSketch™ and ChemDraw® embedded data can be recovered.

To recover chemical data from your Microsoft Office file, please follow these instructions:

1. In ChemDoodle, select the **File>Recover from MS Office File...** menu item. A window will appear as shown here.

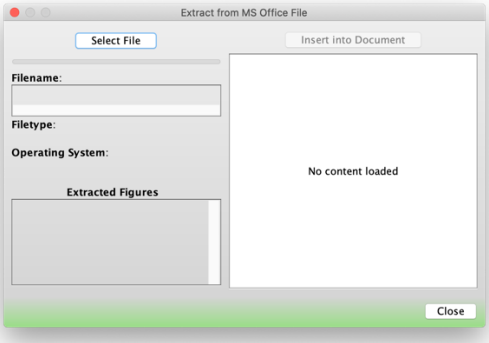


Figure 4: Window to select a Microsoft Office file to recover chemical data from

2. Press the **Select File** button on the top left, and a file chooser will appear. Select the Microsoft Office document you wish to scan for chemical data and press the **Open** button.
3. ChemDoodle will scan the file for compatible chemical data. A progress bar will be animated while the scan is in progress, as the process may take a few seconds or longer. Once complete, the window will display which type of Microsoft Office file was detected and which operating system ChemDoodle believes the file was created on. Any extracted figures will be listed in the **Extracted Figures** list, denoted by the number of molecules and shapes contained in the figure as well as type of chemical data the figure was represented in surrounded by parentheses. This is shown in the next image.

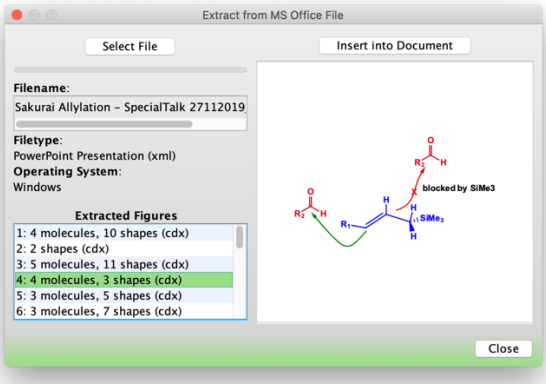


Figure 5: Recovered chemical figures from a Microsoft Powerpoint file

4. The first extracted figure will automatically be selected and displayed in the preview on the right side of the window. Go through the **Extracted Figures** list and select the figure you wish to extract for further editing or any other purpose. You can zoom in on the preview by moving your mouse pointer over the preview panel.
5. Press the **Insert into Document** button to load the extracted figure into your current document.

You can then repaste your content back into the original Microsoft Office file to replace the original figure, but keep in mind the original data will be replaced by ChemDoodle embedded data, and programs other than ChemDoodle may not be able to edit the newly embedded data.

Now, you can share you Microsoft Office files including embedded chemical data with your colleagues on Windows, macOS or even Linux!

Summary

In summary, you can use ChemDoodle to recover chemical data created in ChemDoodle and other chemical programs. This means you will be able to recover data in files you thought were lost long ago, you can edit chemical figures from Microsoft Office files even if you do not have access to Microsoft Office, and you can finally collaborate with chemists in creating documents even if they use a different operating system. When you have ChemDoodle with you, you are in control of your chemical data.

ChemDoodle is a registered trademark of iChemLabs, LLC., ChemDraw is a registered trademark of PerkinElmer, Inc., Microsoft is a registered trademark of Microsoft Corporation, ChemSketch is a trademark of ACD/Labs.

News from the Catalyst Science Discovery Centre & Museum Trust

Contribution from RSC-CICAG Treasurer Dr Diana Leitch MBE, FRSC, email: diana.leitch@googlemail.com

Like many museums and Science Discovery Centres in the UK, Catalyst, which is the only museum of chemistry and the chemical industry in the UK and where I am the Chair of Trustees, reopened with caution after 15 months of closure on 27th May 2021. We have had to work to COVID regulations and our business plan for the next 12 months is worked out on a 20% occupancy rate compared to previous years. Visitors are returning slowly but we don't expect our school visits and uniformed groups' sleepovers to really happen

until the autumn as new waves of COVID strains appear in the north west of England. We have new offerings like Science Birthday Parties and Heritage Afternoon teas in our 4th floor Observatory to attract visitors. Our very successful programme working with MIND (Mental Health charity) continues using science and particularly the chemical sciences as therapy in helping people with mental health challenges. For this work Catalyst won two awards at the recent Chemical Industries Association (CIA) Annual meetings. Catalyst also won the 'Charity of the Year Award' and 'The Most Prestigious Award' too.

Above: Catalyst entrance hall and reception; above right: The ground floor gallery

We have been enabled to survive the long closure period as a result of getting two COVID Recovery Fund (CRF) grants from the Arts Council England (ACE) on the back of being an Accredited Museum. This money should see us through to March 2022 but we are working very hard to get in revenue from admissions, café and shop sales, and more sponsorship from industrial companies, professional societies and individual sponsors so we can continue our work beyond that date. While we were closed the CEO was able to work behind the scenes implementing all the physical changes to our entrance area, café, shop and scientific interactive gallery which were being funded by Wellcome/UKRI so that our overall offering is refreshed and different. We also have a new Gossage Room where visitors can consult our archives and a Henry and John Brunner Room with an associated Brunner exhibition. The Gossages and the Brunner families were hugely important founders of the great chemical industry in the north west in Victorian times. Their greatgrandsons, Peter Gossage and Sir Hugo Brunner, have become very supportive Patrons of Catalyst. Our Education staff have been doing science workshops remotely with uniformed groups including some in Scotland who have been working towards science badges.

I have continued with my adult talks on various science and heritage topics using Zoom and recently have been asked to give a special talk to the AGM of the Cheshire Local History Association on 30th October 2021 on Women Pioneers born in Cheshire. The organisers weren't sure there were any! My talk is entitled '*Two Women Pioneers of the Chemical Information World*' and will feature the redoubtable Miss Angela Haygarth Jackson OBE, the first ever female President of IIS, Head of Information Services at ICI Pharmaceuticals Division when women were rarely senior managers in the chemical industry, and myself. I would have featured Wendy Warr but she was born in Staffordshire! If any of you have special memories of Angela and her activities do please let me know.

News from CAS

Contribution from CICAG member and CAS Custom Services Solutions Engineer Yvonne Pope, email: ypope@acs-i.org

What's New with CAS SciFinder^n

In case you missed it, biosequence searching is now fully implemented within CAS SciFinder^n. BLAST, CDR and Motif searching are now available to mine one of the largest, most comprehensive collections of protein and nucleic acid sequences as well as modified peptides and small molecules. In addition to public datasets and extensive datasets from global patents, the CAS biosequences collection includes millions of human curated and indexed sequences from non-patent literature not found in other databases. CAS has also added Bioscape Analysis to CAS SciFinder^n, allowing the user to visualise the similarity and patent landscape for a set of sequence results. Similarly, Chemscape Analysis visualises the similarity and patent landscape for a set of substance results. When viewing a substance result set that is based on a structure search, an option to 'Create Chemscape Analysis' displays.

There have been many other changes to CAS SciFinder^n in recent months including enhancements to CAS retrosynthesis, a redesigned layout of CAS Commercial Sources information, and usability improvements to provide additional sort and export options for your results.
Please log into CAS SciFinder^n for further details (click on "What's New?" under Account) and to familiarise yourself with these new enhancements.

What's New with STNext®

Patent Status Indicators (STI) for patents and utility models appear in the new PSPI table in CA^{SM}/CAplus^{SM} records. This new information complements your existing legal status workflow and will help you be more efficient in determining which patents require further investigation.

The structure-searchable database, ReaxysFileSub, and its companion bibliographic database, ReaxysFileBib, are now available on STNext. These databases will be updated multiple times each week.

The latest version of the Emtree® thesaurus launched in Embase on STN on May 8, 2021. Emtree remains a great resource for the latest terminology in biomedicine, pharmacology and medical devices and biomedical and drug terminology for searching.

STNext is now also part of the STN IP Protection Suite™. From new users to expert searchers, at all stages of your innovation journey, CAS has the solution to meet your scientific IP needs.

- STNext - The premier IP information platform and the choice of patent experts
- CAS Scientific Patent Explorer™ - a specialised and easy-to-use solution for scientists and IP searchers
- CAS Search Guard™ - expanded capacity and trusted experience when you need it most
- FIZ PatMon - the easy-to-use Patent Monitoring service for everyone engaged in the patent life cycle

To learn more about the STN IP Protection Suite, please visit cas.org or reach out to your CAS representative.

Upcoming Webinars

The extremely popular CAS SciFinder^n public webinar series continues to attract thousands of regular attendees. Upcoming seminars include:
Advanced Reference and Substance Searching in CAS SciFinder^n, June 23, 2021
Retrosynthesis Enhancements in CAS SciFinder^n, August 18, 2021
To register for these sessions and view recordings of previous events, please visit our Events Page.

Predicting New Chemistry: Impact of High-Quality Training Data on Prediction of Reaction Outcomes

In a recent study, scientists at Bayer demonstrated the significant impact that scientist-curated reactions from the CAS Content Collection have on the predictive power of a synthesis planning model. Accuracy in prediction of outcomes in rare reaction classes increased significantly – a boost of 32 percentage points – expanding understanding into new, useful chemistry. A white paper summarising the work can be downloaded from the CAS website.

Training set / Test set	V1-base	V2-cas
CAS reaction data only	16%	48%
Negative reaction data only	57%	57%

Table 2: Accuracy for supplemental data only

To discuss how CAS can help you maximize your individual scientific information advantage, please reach out to the CAS Custom Services Team.

CAS Common Chemistry™ expands collection of publicly available chemical information
To strengthen the accuracy of publicly available scientific information, CAS has expanded the CAS Common Chemistry resource that now provides authoritative information on nearly 500,000 substances from CAS REGISTRY®. The collection represents substances commonly found in consumer products, on regulatory lists and as part of introductory chemistry curricula. Enhanced search functionality and an application programming interface (API) enable interested individuals to quickly find chemical names, CAS Registry Numbers®, structures or basic compound properties.

First launched in 2009, CAS Common Chemistry initially provided public access to several thousand substances and associated data. In collaboration with thought leaders in the chemistry and open science community, CAS scientists dramatically expanded the community resource with a focused subset of substances from CAS REGISTRY. Optimised for mobile devices, CAS Common Chemistry now makes accurate chemical information accessible from anywhere.

In addition to expansion of included substances, enhanced capabilities will meet evolving demands for information. Newly added API capabilities support the growing needs of digital workflows and cheminformatics initiatives with direct lookup functionality. Reusable licensing of content is ensured through a Creative Commons Attribution-NonCommercial (CC BY-NC 4.0) license.

Chemical Information / Cheminformatics and Related Books

Contributed by RSC CICAG Newsletter Editor Stuart Newbold, email: stuart@psandim.com

Computational Chemistry from A to Z: A Concise Encyclopedia

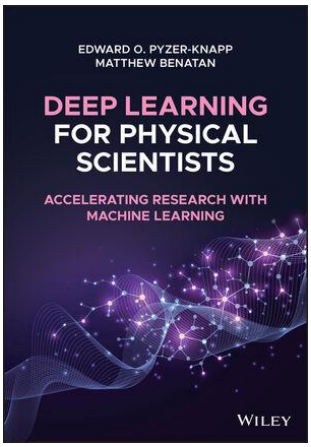
This first mini encyclopaedia devoted to computational chemistry contains over 1,000 alphabetical entries, covering basic as well as specialised terms and concepts in a precise and concise way, including:

- mathematical foundations
- wave function- based methods
- density functional theory
- relativistic quantum mechanics
- semi-empirical and force-field methods
- population analysis tools
- cheminformatics

A quick and handy reference.

Wiley
Michael Bühl, Maciej Gutowski, John B. O. Mitchell, Markus Reiher

Deep Learning for Physical Scientists: Accelerating Research with Machine Learning



Deep Learning for Physical Scientists: Accelerating Research with Machine Learning delivers an insightful analysis of the transformative techniques being used in deep learning within the physical sciences. The book offers readers the ability to understand, select, and apply the best deep learning techniques for their individual research problem and interpret the outcome.

Designed to teach researchers to think in useful new ways about how to achieve results in their research, the book provides scientists with new avenues to attack problems and avoid common pitfalls and problems. Practical case studies and problems are presented, giving readers an opportunity to put what they have learned into practice, with exemplar coding approaches provided to assist the reader.

From modelling basics to feed-forward networks, the book offers a broad cross-section of machine learning techniques to improve physical science research. Readers will also enjoy:

- A thorough introduction to the basic classification and regression with perceptrons
- An exploration of training algorithms, including back propagation and stochastic gradient descent and the parallelisation of training
- An examination of multi-layer perceptrons for learning from descriptors and de-noising data
- Discussions of recurrent neural networks for learning from sequences and convolutional neural networks for learning from images
- A treatment of Bayesian optimisation for tuning deep learning architectures

Perfect for academic and industrial research professionals in the physical sciences, Deep Learning for Physical Scientists: Accelerating Research with Machine Learning will also earn a place in the libraries of industrial researchers who have access to large amounts of data but have yet to learn the techniques to fully exploit that access.

Wiley
Edward O. Pyzer-Knapp & Matthew Benatan

News from AI3SD

Contribution from AI3SD Network+ Coordinator Dr Samantha Kanza, email: s.kanza@ai3sd.org

AI3SD Intern Cohort 2021
AI3SD are delighted to announce their internship cohort of 2021. You can read all about our interns here. The projects they are working on are interdisciplinary and include both cutting-edge AI and cutting-edge chemical discovery and demonstrate how and why they are relevant to AI3SD. The projects must be able to demonstrate valuable outputs both with respect to developing student skills and providing impact to AI3SD.

AI3SD & PSDS Skills4Scientists Series 2021
As part of the work on our internship programme, we are running a Skills4Scientists series. This series has been organised as a joint venture between the Artificial Intelligence for Scientific Discovery Network+ (AI3SD) and the Physical Sciences Data-Science Service (PSDS). This series will run over the summer and aims to educate and improve scientists' skills in a range of areas including research data management, python, version control, ethics, and career development. The series is primarily aimed at final year undergraduates / early-stage PhD students. Although of course anyone interested in any of the topics is welcome to join! The series starts on 7th July 2021 and runs every other Wednesday and Thursday afternoon until the end of August. Each seminar session will commence at 14:00 and typically finish around 17:00 UK time. The final event of the series will be 2-day poster symposia where attendees will be invited to present their research to a range of experts from industry and academia. The full details of each event can be found here.

AI3SD & DA ML4MC Summer School 2021
We are pleased to announce that AI3SD have teamed up with the Directed Assembly Network to run a virtual summer school this summer on Machine Learning 4 Materials & Chemicals which encompass our overlapping Network interests of AI, Machine Learning, Artificial Photosynthesis, Biomimetic Materials, Crystal Design & Engineering, Materials, Molecules, Photochemistry, Photocatalysis and Supramolecular Chemistry. This virtual summer school will run between the 6th July and the 24th August, and will run on British Summer Time (BST). Entries for this summer school have now closed, but we will be making the videos of the talks given as part of this summer school available on our AI3SD YouTube Channel so keep your eyes peeled for those.

AI3SD Autumn Seminar Series 2021
We are currently organising our Autumn Seminar Series which will begin in October. If you are interested in participating in this or would like to express interest about us organising events on specific topics, please get in touch with Dr Samantha Kanza on s.kanza@soton.ac.uk.

Other Chemical Information Related News

Contribution from RSC CICAG Newsletter Editor Stuart Newbold

[*All hyperlinks correct & working as of 6 Jul 2021*]

New RSC Journal

Digital Discovery, is a new open access journal from the Royal Society of Chemistry that will chart the discovery of new chemicals, processes and biological systems through new tech development. The journal covers chemistry, biology, physics, and materials & biomedical sciences, and complements the expanded scope of PCCP. A May 2021 interview with Editor in Chief, Alán Aspuru-Guzik, explains more.
Source: RSC Publishing

AI Firm Insilico closes a $255M Financing round as it builds on Discovery

Research

Insilico Medicine, a developer of artificial intelligence technology for drug discovery, has raised $255 million in series C financing. The investment round is the second this year to top $200 million for an AI-based drug discovery firm. The investment follows Insilico's announcement in February that it used AI to identify a novel target and develop an investigational small-molecule therapy to treat idiopathic pulmonary fibrosis. Claiming a first, Insilico says it identified the target, developed molecules including selecting a lead at a cost of $1.5m, and completed preclinical trials in 18 months.

https://cen.acs.org/business/investment/AI-firm-Insilico-closes-255/99/web/2021/06

Source: Chemical & Engineering News

ACS Journals continue to be the most cited in Chemistry

ACS journals reached new levels of impact, citations and output in 2020. The 2021 Web of Science™ Journal Citation Reports™ from Clarivate, detail the extent to which ACS journals influence and shape chemistry across its many fields. Throughout their entire portfolio, ACS journals achieved a median 2020 Journal Impact Factor™, Journal Citation Reports™ (Clarivate, 2021), of 5.097, a 14% increase from 2019. A Journal Impact Factor™, calculated yearly by Clarivate, provides a measure of how often an average article from a journal has been cited in the previous year. This increase demonstrates that articles published in ACS journals are growing ever more relevant to scientists. Total citations also grew by over 550,000, a more than 15% increase, and citations grew from the previous year for every one of ACS' journals.

https://www.knowledgespeak.com/news/american-chemical-society-journals-continue-to-be-the-most-cited-in-chemistry/

Source: Knowledgespeak

Clarivate to Acquire ProQuest

Clarivate, a global leader in providing trusted information and insights to accelerate the pace of innovation, has announced a definitive agreement to acquire ProQuest, a leading global software, data and analytics provider to academic, research and national institutions, from Cambridge Information Group, a family-owned investment firm, and other partners including Atairos, for $5.3 billion, including refinancing of ProQuest debt. The consideration for the acquisition is approximately $4.0 billion in cash and $1.3 billion of equity. The transaction, which is subject to customary closing conditions, including regulatory approvals, is expected to close during the third quarter of 2021.

https://about.proquest.com/en/news/2021/clarivate-to-acquire-proquest/

Source: Proquest

The Royal Society sets 75% threshold to 'flip' Research Journals to OA in the next five years

The Royal Society, in an exciting new chapter for its scientific publishing, sets out how it will transition its primary research journals to open access and make more of its world-leading research available to all. Following a review by its Council, the Royal Society has committed to 'flipping' the journals Biology Letters, Interface, Proceedings A, and Proceedings B to a fully open access model when 75% of articles are being published open access. This transition will be driven chiefly by the expansion of Read & Publish agreements with major research institutions, enabling their scientific research output to be published open access in the Society's journals.

https://www.knowledgespeak.com/news/the-royal-society-sets-75-threshold-to-flip-research-journals-to-oa-in-the-next-five-years/

Source: Knowledgespeak

Ping An makes Breakthrough in Artificial Intelligence-Driven Drug Research

Research by Ping An Healthcare Technology Research Institute and Tsinghua University has led to a promising deep learning framework for drug discovery. The findings were published in "*An effective self-supervised framework for learning expressive molecular global representations to drug discovery*" in Briefings in Bioinformatics, a peer-reviewed bioinformatics journal. It marks a major technology breakthrough for the Group in the field of AI-driven pharmaceutical research.

https://www.prnewswire.co.uk/news-releases/ping-an-makes-breakthrough-in-artificial-intelligence-driven-drug-research-804496316.html
Source: Cision PR Newswire

CAS expands CAS Common Chemistry resource; further strengthens accuracy of Publicly Available Scientific Information
CAS has expanded the CAS Common Chemistry resource. To strengthen the accuracy of publicly available scientific information, CAS Common Chemistry now provides authoritative information on nearly 500,000 substances from CAS REGISTRY®. The collection represents substances commonly found in consumer products, on regulatory lists and as part of introductory chemistry curricula. As part of ACS, CAS is committed to providing the global community with accurate chemical information. Aligned with the ACS Open Science initiative, CAS Common Chemistry broadens access for everyone who seeks reliable scientific information on routinely encountered chemicals. Enhanced search functionality and an application programming interface (API) enable interested individuals to quickly find chemical names, CAS Registry Numbers®, structures or basic compound properties.
https://www.knowledgespeak.com/news/cas-expands-cas-common-chemistry-resource-further-strengthens-accuracy-of-publicly-available-scientific-information/
Source: Knowledgespeak

Clarivate announces 2021 update to its annual Journal Citation Reports
Clarivate Plc, a global leader in providing trusted information and insights to accelerate the pace of innovation, has released the 2021 update to its annual Journal Citation Reports (JCR)™. The annual JCR release enables the research community to evaluate the world's high-quality academic journals using a range of indicators, descriptive data and visualisations. The reports are used extensively by academic publishers across the globe to evaluate the impact of their journals relative to their field and promote them to the research community.
https://www.knowledgespeak.com/news/clarivate-announces-2021-update-to-its-annual-journal-citation-reports/
Source: Knowledgespeak

Atypon Acquires RedLink®, the Data and Engagement home for Scholarly Communications
Atypon, the technology leader in the scholarly publishing market, has acquired RedLink®, the provider of key data and insights for academic publishers, libraries, and consortia. RedLink's portfolio includes Remarq®, the popular article-sharing, commenting, and annotation platform for researchers.
https://redlink.com/atypon-acquires-redlink-the-data-and-engagement-home-for-scholarly-communications/
Source: Redlink

The Emerging Technologies Shaping 2021 Best Practices for Information Professionals
Buzzwords like artificial intelligence and machine learning are used by nearly every vendor in 2021. These emerging technologies present exciting new opportunities for information professionals to help researchers turn scientific content and data into insights to accelerate innovation, but the specifics are often lacking.
http://www.copyright.com/blog/emerging-technologies-2021-best-practices-information-professionals/
Source: CCC

Cambridge University Press announces new Open Access pilot scheme - Flip it Open
A new Open Access pilot scheme from Cambridge University Press will turn conventional publishing models on their head by making academic monographs that sell the most copies available online for free
https://www.knowledgespeak.com/news/cambridge-university-press-announces-new-open-access-pilot-scheme-flip-it-open/
Source: Knowledgespeak

Taylor & Francis strengthens its partnership with the Research Data Alliance (RDA)
Taylor & Francis has joined the RDA as an organisational member, strengthening ties with a leading data organisation that is focused on facilitating data reuse and open data sharing. RDA is a multi-stakeholder

community who work collaboratively and creatively together to overcome barriers to data sharing. The key players in the RDA community are researchers, librarians, data professionals, repositories, infrastructure groups and publishers. Taylor & Francis colleagues have been involved in RDA working groups for many years, chairing or participating working groups on data sharing policies, repository guidance and handling specific subject-based data. This enhanced partnership highlights our shared goal of supporting researchers to openly share data in order to improve discoverability, reproducibility and quality of research.
https://www.stm-publishing.com/taylor-francis-strengthens-its-partnership-with-the-research-data-alliance-rda/
Source: STM Publishing News

Springer Nature Advances its Machine-Generated Tools and offers a new Book Format with AI-based Literature Overviews
Climate, Planetary and Evolutionary Sciences: A Machine-Generated Literature Overview marks a further step in the innovation journey of publishing services at Springer Nature.
Following the publication of its first machine-generated book in 2019, Springer Nature has now deployed its AI expertise to create a new publication format which focuses on literature reviews. While the first book on lithium-ion batteries was entirely AI-based, this new format takes an innovative hybrid approach of blending human-machine interaction. The new product is a mixture of human-written text and machine-generated literature overviews, which sees an author putting these machine-generated reviews, created from a large set of previously published articles in Springer Nature journals, into book chapters and providing a scientific perspective.
https://group.springernature.com/gp/group/media/press-releases/advances-its-machine-generated-tools-with-ai-based-lit-overviews/19129322
Source: Springer Nature

New Database expands access to Open Content on OCLC FirstSearch Service
A new database of open access content on the OCLC FirstSearch service will help researchers easily find open access resources from familiar content providers. Developed in response to requests from FirstSearch users, the Open Access Content database expands on OCLC's 20-year history of partnering with the world's leading content providers to represent high-quality open content in WorldCat and OCLC services.
https://www.stm-publishing.com/new-database-expands-access-to-open-content-on-oclc-firstsearch-service/
Source: STM Publishing News

Linked Knowledge: A Data-Driven Approach to Life Sciences
In any knowledge area, content of interest can be found across many different data sources. We can think of a data source as a "bucket" of information items, where each item belongs to a particular category or type, according to the nature of the information it contains – journal articles, patents, trials/studies, chemical substances, etc. A bucket might contain one or several data categories and potentially thousands of items, which are often semantically related to but disconnected from items in other buckets. That is why we usually speak of distinct data sources as isolated "information silos."
http://www.copyright.com/blog/linked-knowledge-a-data-driven-approach-to-life-sciences/
Source: CCC

IET announces launch of Inspec Analytics Plus to enhance analysis of Collaboration and Research Impact
The IET has launched Inspec Analytics Plus, expanding the capability of the research intelligence tool Inspec Analytics to support citation and collaboration analysis. Inspec Analytics Plus incorporates citation data and enhanced collaboration analysis at a very granular level to uncover deeper insights into the impact of research. This level of precision analytics allows users to further explore global research, and assess the value and impact of work, by augmenting Inspec Analytics with new data and functionality.
https://www.stm-publishing.com/iet-announces-launch-of-inspec-analytics-plus-to-enhance-analysis-of-collaboration-and-research-impact/
Source: STM Publishing News

Accessing British Library Online Content and Services
The British Library is planning a return to public spaces that keeps everyone safe. Meanwhile the library is always open online and has published a guide to help online visitors explore their collection.
https://www.bl.uk/news/2020/march/access-to-the-british-library-during-temporary-closure
Source: British Library

Elsevier's Mini Program Launched on WeChat Brings Quality Editing Straight to your Smartphone
Academics can now access language editing and translation services on their smartphone with the launch of Elsevier's Mini Program on WeChat. The service provides convenient and easy access to get quotes and submit inquiries with just a few finger taps. It can be used on the go and the final result is a first-rate manuscript completely free of language errors.
https://www.stm-publishing.com/elseviers-mini-program-launched-on-wechat-brings-quality-editing-straight-to-your-smartphone/
Source: STM Publishing News

Artificial Chemist 2.0
A group of researchers from North Carolina State University and the University at Buffalo have developed an intelligent robotic material synthesiser to accelerate the development and manufacturing of semiconductor nanomaterials for applications in next-generation photonic devices. This is the second generation of the Artificial Chemist technology, that is, a fluidic robo-chemist operated with an artificial intelligence (AI)-guided decision-making agent.
https://www.advancedsciencenews.com/artificial-chemist-2-0/
Source: Advanced Science News

Remote Education inspires new ways to teach Chemistry Lab Courses
The pandemic forced chemistry instructors to rethink how to teach the lab component of undergraduate chemistry classes.
https://cen.acs.org/education/undergraduate-education/Remote-education-inspires-new-ways-to-teach-chemistry-lab-courses/99/i24
Source: C&EN

PLOS Adopts Copyright Clearance Center's RightsLink® for Scientific Communications to Manage New Community Action Publishing Model
The Public Library of Science has announced it is using RightsLink® for Scientific Communications (RLSC) to manage its Community Action Publishing (CAP) model, which aims to eliminate author APCs in order to make its Open Access (OA) journals truly Open to Read and Open to Publish. PLOS Biology and PLOS Medicine are the first two PLOS journals live on RLSC, the comprehensive scholarly communications workflow solution from Copyright Clearance Center, Inc. (CCC), a leader in advancing copyright, accelerating knowledge, and powering innovation.
https://www.stm-publishing.com/plos-adopts-copyright-clearance-centers-rightslink-for-scientific-communications-to-manage-new-community-action-publishing-model/
Source: STM Publishing News

Empowering Competitive Intelligence Teams with better access to Scientific Literature
Manually sifting through vast amounts of scientific literature and data is a challenge for nearly every area of an R&D intensive organisation. Missed information can be extremely costly, particularly for competitive intelligence teams who are responsible for understanding the landscape that informs their company's go-to-market strategy.
http://www.copyright.com/blog/empowering-competitive-intelligence-teams-better-access-scientific-literature/
Source: CCC

Springer Nature leads drive for Open Access across Europe with latest Transformative Agreement with Spain

Publisher holds the highest number of national transformative agreements in Europe enabling researchers from over 2,100 affiliated institutions to publish OA.

https://group.springernature.com/gp/group/media/press-releases/springer-nature-leads-drive-for-oa--spain/19134258

Source: Springer Nature

The European Commission launches Open Access Publishing Platform for Scientific Papers

The European Commission launched Open Research Europe, a publishing platform for scientific papers that will be accessible to everyone. The platform will present the results of research funded by Horizon Europe, the EU research and innovation programme for 2021-2027, and its predecessor, Horizon 2020.

https://www.infotoday.eu/Articles/News/Featured-News/The-European-Commission-launches-open-access-publishing-platform-for-scientific-papers-146084.aspx

Source: Information Today

The Evolution of Text Mining – Trends we're seeing across R&D Organisations

It's no secret that information available in the digital ecosystem is rapidly expanding. Three million papers are published in scholarly journals annually, and that's just one high value content type for R&D–intensive organisations that may also need quick and easy synthesis of patent, clinical, and other types of content.

http://www.copyright.com/blog/trends-evolution-text-mining/

Source: CCC

Heidi Fraser-Krauss named new CEO of Jisc

Heidi Fraser-Krauss, currently executive director of corporate services at the University of Sheffield, will join Jisc as the new CEO, effective September 16, 2021. She will succeed Paul Feldman, who is retiring after a six-year stint at the helm.

https://www.knowledgespeak.com/news/heidi-fraser-krauss-named-new-ceo-of-jisc/

Source: Knowledgespeak

DSTL Celebrates 20 Remarkable Years

The Defence Science and Technology Laboratory (Dstl) celebrates its 20th anniversary on 2nd July 2021 as the science inside UK defence and security.

https://www.gov.uk/government/news/dstl-celebrates-20-remarkable-years

Source: Gov.uk

China 'Pursuing National Open Science Strategy'

China is working on a master plan for the internationalisation of its domestic journals and plans to pursue an open science strategy at a national level.

https://www.researchinformation.info/news/china-pursuing-national-open-science-strategy

Source: Research Information

SciBite launches SaaS versions of CENtree Ontology Management and TERMite Text Analysis Solutions

Semantic technology company SciBite, an Elsevier company, has unveiled its new Software-as-a-Service (SaaS) version of TERMite, SciBite's named entity recognition engine, and its CENtree™ ontology management platform. Researchers can use the new cloud-hosted TERMite and CENtree as part of their digital workplace, offering greater flexibility to enable them to work from multiple locations and settings.

https://www.knowledgespeak.com/news/scibite-launches-saas-versions-of-centree-ontology-management-and-termite-text-analysis-solutions/

Source: Knowledgespeak

Artificial Intelligence, the EU, Libraries and You

The European Union have proposed new rules and actions regarding Artificial Intelligence, in line with its desire to make "Europe fit for the Digital Age". Although most commentary regarding the proposed AI regulations have centred on "Big Tech", the library community is also affected.

https://www.infotoday.eu/Articles/News/Featured-News/Artificial-Intelligence-the-EU-Libraries-and-You-147128.aspx

Source: Information Today

An upcoming RLUK Report looks at the role of Libraries in Scholarly Research

The role of academic and research libraries as active participants and leaders in the production of scholarly research report was commissioned by RLUK, in partnership with the AHRC, in order to investigate the role of academic and research library staff in the initiation, production, and dissemination of academic and scholarly research. The research project has been led and delivered by Evidence Base, a research consultancy within Birmingham City University, and their associate

https://www.knowledgespeak.com/news/an-upcoming-rluk-report-looks-at-the-role-of-libraries-in-scholarly-research/

Source: Knowledgespeak

AI and Data Science Expansion for DSTL in the North

The Defence Science and Technology Laboratory (Dstl) is creating a new unit based at the National Innovation Centre for Data (NICD) in Newcastle.

https://www.gov.uk/government/news/ai-and-data-science-expansion-for-dstl-in-the-north

Source: Gov.uk

Science Europe launches new Strategy Plan 2021-2026 and Multi-annual Action Plan 2021-2026

Science Europe recently launched its new Strategy Plan 2021-2026 that outlines how the Brussels-based association of research funders and performers will drive the development of European Research Area (ERA) policies and build the strongest possible research ecosystem for the benefit of science, researchers, and society. The publication lays out an updated vision, mission, values, and set of strategic priorities.

https://www.knowledgespeak.com/news/science-europe-launches-new-strategy-plan-2021-2026-and-multi-annual-action-plan-2021-2026/

Source: Knowledgespeak

F1000 working on 'Digital Twin' Platform Launches

F1000 is collaborating with two Chinese customers to develop open research publishing platforms dedicated to the research and application of collaborative robots and 'digital twin' technologies. Both will be the world's first open publishing platforms in their fields and will launch for submission in July 2021. The platforms will utilise F1000's open research publishing model, enabling all research outputs to be published open access, as well as combining the benefits of pre-printing (providing rapid publication with no editorial bias) with mechanisms to assure quality and transparency (invited and open peer review, archiving and indexing). They also offer researchers an open and transparent peer review process and have a mandatory FAIR data policy to provide full and easy access to the source data underlying the results.

https://www.researchinformation.info/news/f1000-working-digital-twin-platform-launches

Source: Research Information

SLAS partners with Elsevier to Transition two of its Journals to Gold Open Access Publishing Model

The Society for Laboratory Automation and Screening (SLAS) has announced the transition of SLAS Discovery and SLAS Technology to a Gold Open Access publishing model as of January 1, 2022. SLAS will partner with Elsevier, which will make both journals available through its online platform, ScienceDirect, in turn, making articles instantly and easily accessible to scientific audiences worldwide.

https://www.knowledgespeak.com/news/slas-partners-with-elsevier-to-transition-two-of-its-journals-to-gold-open-access-publishing-model-effective-january-2022/
Source: Knowledgespeak

Sosei Heptares and InveniAI enter a Multi-target AI-powered and GPCR-focused Drug Discovery Collaboration
Creating transformative therapeutics for immune diseases by combining InveniAI's validated AI-based platform, AlphaMeld®, with Sosei Heptares' GPCR Structure-based Drug Design and early development capabilities.
https://www.prnewswire.co.uk/news-releases/sosei-heptares-and-inveniai-enter-a-multi-target-ai-powered-and-gpcr-focused-drug-discovery-collaboration-820787038.html
Source: Cision PR Newswire

Springer Nature launches new Publication Portal for Dissertations
Researchers wishing to publish their dissertation or postdoctoral thesis with Springer Nature can now benefit from a new streamlined and author-friendly publication portal. BookSubMarine (Submitted book Manuscripts are in evaluation) supports authors of theses in STM (Science Technology Medicine) and HSS (Humanities Social Sciences) throughout the entire publication process. Prospective authors submit the required data to the portal, such as their manuscript and a review, and select the desired publication model (Open Access, Basic or Premium).
https://www.stm-publishing.com/springer-nature-launches-new-publication-portal-for-dissertations/
Source: STM Publishing News

DeepDyve Digital Library offers affordable access to research literature
DeepDyve announced the availability of the DeepDyve Digital Library, a fully integrated platform aimed at helping research organisations discover, access and manage research papers. The Digital Library is a one-stop platform designed for small to medium-sized teams and organisations needing easy and affordable solutions for reading and organising scientific papers. Although the press release states that this product doesn't need information professionals or IT staff, it should be on the radar screen of librarians, particularly those in smaller, poorly-funded institutions and companies.
https://www.infotoday.eu/Articles/News/Featured-News/DeepDyve-Digital-Library-offers-affordable-access-to-research-literature-146508.aspx
Source: Information Today

Clarivate Releases New Visualisation Tool for Multi-Dimensional Individual Researcher Performance
Clarivate Plc has released a new interactive tool that showcases the range of a researcher's publication and citation impact in a single visualisation. The Web of Science™ Author Impact Beamplots will enable users of research metrics to engage more actively with data and provide new opportunities to conduct research assessment in a responsible manner.
https://www.stm-publishing.com/clarivate-releases-new-visualization-tool-for-multi-dimensional-individual-researcher-performance/
Source: STM Publishing News

Dotmatics Collaboration to provide a Single Software Interface for Lab Digitalisation
Dotmatics, a scientific informatics software and services company has announced a partnership with HighRes Biosolutions (HighRes), that will enable scientists to plan experiments and run instruments from the ELN. The collaboration combines the high-throughput laboratory automation capability of Dotmatics' electronic laboratory notebook (ELN) with HighRes' Cellario instrument control software enabling R&D scientists to plan experiments, run individual instruments, and publish and analyse data within a single software interface. The integration will save researchers' time and improve ease of operation as a further step towards the Lab of the Future (LoTF).
https://www.scientific-computing.com/news/dotmatics-collaboration-provide-single-software-interface-lab-digitalisation

Source: Scientific Computing World

EBSCO Information Services unveils the EBSCO eBooks™ Open Access Monograph Collection
EBSCO Information Services (EBSCO) is releasing a new e-book collection containing thousands of high-quality open access (OA) e-books from the world's most trusted university presses and scholarly publishers. The EBSCO eBooks™ Open Access Monograph Collection is DRM-free and available at no cost on the EBSCOhost® platform and easily discoverable via EBSCO Discovery Service.
https://www.knowledgespeak.com/news/ebsco-information-services-unveils-the-ebsco-ebooks-open-access-monograph-collection/
Source: Knowledgespeak

ResearchGate and Wiley Deepen Partnership with Content Pilot to Deliver New Value for Researchers
ResearchGate and Wiley today announced a pilot that will make articles published in selected Wiley journals available on the ResearchGate platform. The pilot aims to save authors time, enhance the visibility and discoverability of their peer-reviewed research, and measure the impact of their work. Wiley and ResearchGate enter this pilot, which builds on their partnership agreement announced in May 2020, to better meet the needs of researchers by encouraging greater discussion and collaboration around timely publications.
https://www.stm-publishing.com/researchgate-and-wiley-deepen-partnership-with-content-pilot-to-deliver-new-value-for-researchers/
Source: STM Publishing News

ACS announces transformative 'read and publish' deal with the Conference of Rectors of Spanish Universities and the Spanish National Research Council
The Publications Division of the ACS has announced a landmark transformative "read and publish" agreement with the Conference of Rectors of Spanish Universities (Crue) and the Spanish National Research Council (CSIC). Under the deal, researchers affiliated with institutions of Crue and the CSIC will now be able to publish their research articles open access with no additional costs.
https://www.knowledgespeak.com/news/acs-announces-transformative-read-and-publish-deal-with-the-conference-of-rectors-of-spanish-universities-and-the-spanish-national-research-council/
Source: Knowledgespeak

Leading Patent Data Platform IFI CLAIMS joins Digital Science
Digital Science, a technology company serving stakeholders across the research ecosystem, has announced that it has fully acquired IFI CLAIMS, a leading international platform for patent and innovation data. Digital Science originally invested in IFI CLAIMS in 2017 as part of its aim to diversify the search and analytics landscape by bringing together awarded grants data, publication and citation data with patent data to allow better contextualisation of research – from funding through publication to translation and impact. IFI CLAIMS provides 137 million patents to Dimensions and is the data source behind the patent facet in the Altmetric donut.
https://www.digital-science.com/press-release/leading-patent-data-platform-joins-digital-science/
Source: Digital Science

ResoluteAI unveils Resolute Research Network, a Database of Key Opinion Leaders in STEM
ResoluteAI, the research platform for science, has announced the release of the Resolute Research Network (RRN), a comprehensive global database of Key Opinion Leaders (KOLs) in Science, Technology, Engineering, and Mathematics (STEM). RRN aggregates the research interests, biographical information, awards, and publication information of over 230,000 academic faculty from more than 350 universities and research institutes around the world. The database is searchable and can highlight connections between faculty and their patents, investigatory roles in clinical trials, and co-authors.
https://www.knowledgespeak.com/news/resoluteai-unveils-resolute-research-network-a-database-of-key-opinion-leaders-in-stem/
Source: Knowledgespeak

Wiley announces Hindawi Acquisition

John Wiley & Sons has announced the acquisition of Hindawi Limited, for a total purchase price of $298 million. Wiley says the purchase adds 'quality, scale and growth to the company's open access publishing program'. Open access is a rapidly growing scholarly publishing model that allows peer-reviewed articles to be read and shared immediately, making important research broadly available. Hindawi has played a critical role in advancing gold open access, an OA model in which validated articles are made immediately available for reading and re-use following the payment of a publication fee.

https://www.researchinformation.info/news/wiley-announces-hindawi-acquisition

Source: Research Information

University of Oxford set to launch Pandemic Sciences Centre

The Pandemic Sciences Centre, which will include a number of core institutes, will harness the strong global research collaborations that the University of Oxford has developed over more than forty years. Its mission will be to ensure that the world is better equipped to create global, and equitable science-driven solutions to prepare for, identify, and counter future pandemic threats. The new centre will also build on the strong and unique collaborations developed in record time across national borders between academia, industry and public health bodies during the coronavirus pandemic.

https://www.knowledgespeak.com/news/university-of-oxford-set-to-launch-pandemic-sciences-centre/

Source: Knowledgespeak

ChemRxiv™ Enhances User Experience

ChemRxiv, the leading preprint server for the chemical sciences operated by the ACS, Chinese Chemical Society, Chemical Society of Japan, German Chemical Society (GDCh), and the RSC, has stated that the server will be hosted on the Cambridge Open Engage platform beginning in the second quarter of 2021.

https://www.chemistryviews.org/details/news/11290551/ChemRxiv_Enhances_User_Experience.html

Source: ChemistryViews

RAMBO Speeds Searches on huge DNA Databases

Method cuts indexing times from weeks to hours, search times from hours to minutes. Computer scientists are sending RAMBO to rescue genomic researchers who sometimes wait days or weeks for search results from enormous DNA databases.

https://www.sciencedaily.com/releases/2021/06/210628152920.htm

Source: ScienceDaily

Atos launches 'ThinkAI'

Atos has launched ThinkAI, its secure end-to-end scalable offering which enables organisations to successfully design, develop, and deliver high-performance AI applications. ThinkAI is for organisations using traditional high-performance computing that want to run more accurate and faster simulations thanks to AI applications, and also for those developing AI applications that need more computing power.

https://www.scientific-computing.com/news/atos-launches-thinkai

Source: Scientific Computing World

CAS unveils new Brand Expanding Solutions and Capabilities Addressing rising need for Curated Scientific Data and expertise to Accelerate Discovery

CAS has launched a new brand reflecting the broadening scope of its solutions and capabilities critical to advancing scientific discovery. This new brand mirrors the organisation's on-going evolution to enable new and smarter uses of scientific information as a partner to R&D leaders pursuing more efficient, collaborative, and customised innovation models. This launch is the culmination of global research and consultation with stakeholders across commercial, academic and government segments to distil the most impactful elements that CAS provides as an innovation partner. The new brand embodies the unique role CAS plays in creating connections between discoveries being made around the world to accelerate breakthroughs. The new logo

evolves the molecular icon, whose three colours represent the hindsight, insight and foresight that are foundational to scientific discovery.
https://www.knowledgespeak.com/news/cas-unveils-new-brand-expanding-solutions-and-capabilities-addressing-rising-need-for-curated-scientific-data-and-expertise-to-accelerate-discovery/
Source: Knowledgespeak

67 Bricks strengthens relationship with the Royal Society of Chemistry
Alongside its activities as the professional body for chemists in the UK, the RSC works across industry, policy and education, bringing people together to spark new ideas and new partnerships, and supporting teachers to inspire future generations of scientists. It's through their work as a high-quality scientific publisher that 67 Bricks have worked alongside them for the last five years – this new agreement highlights their commitment to reinvigorating their processes and services so they are fully fit to advance chemistry in the digital age.
https://www.stm-publishing.com/67-bricks-strengthens-relationship-with-the-royal-society-of-chemistry/
Source: STM Publishing News

New Book Publishing concept combines Taylor & Francis Books Heritage with F1000's Open Research Model
Open Plus Books will enable authors or book editors to publish a book open access first on the open research platform within just a few days of submission. Each chapter is a standalone piece, enabling chapters to be published as soon as they are accepted, and additional chapters to be added at any time. Chapters can be updated, revised, and amended individually by the author to reflect changes in research, knowledge, policy, legislation, or standards.
https://newsroom.taylorandfrancisgroup.com/new-book-publishing-concept-combines-taylor-francis-books-heritage-with-f1000s-open-research-model/
Source: Taylor & Francis

Elsevier launches Browser-based QUOSA to bring order to Life Sciences Literature Management
The company has launched a fully browser-based literature management solution, QUOSA, to help life science companies centralise the process of collecting and sharing scientific information while driving copyright compliance.
https://www.elsevier.com/about/press-releases/science-and-technology/elsevier-launches-browser-based-quosa-to-bring-order-to-life-sciences-literature-management
Source: Elsevier

Digital Science acquires Ripeta
Digital Science, a technology company serving stakeholders across the research ecosystem, has fully acquired Ripeta, an AI-based company aiming to help build trust in science. This is a natural development based on Digital Science's previous support and investment in the US-based start-up, which aims to make research more reproducible by identifying and highlighting elements of scholarly manuscripts that either appear to be difficult to reproduce or where additional provenance would increase the trust that can be placed in the paper.
https://www.knowledgespeak.com/news/digital-science-acquires-ripeta/
Source: Knowledgespeak

Researcher launches its free Profile Manager, providing Publishers and Content Owners with a new and free way to reach and engage directly with Researchers
The Researcher Profile Manager is an all-in-one content management, analytics and advertising platform.
https://www.stm-publishing.com/researcher-launches-its-free-profile-manager-providing-publishers-and-content-owners-with-a-new-and-free-way-to-reach-and-engage-directly-with-researchers/
Source: STM Publishing News

AAAS to move Science Journals to new Online Platform to enhance User Experience
The American Association for the Advancement of Science (AAAS), the world's largest general scientific society and publisher of the Science family of journals, will be moving all its online content to Atypon's online

publishing platform, Literatum for a better user experience. With this move, AAAS seeks to make all content across the Science family of journals more integrated, discoverable, and visually compelling for the reader.
https://www.knowledgespeak.com/news/aaas-to-move-science-journals-to-new-online-platform-to-enhance-user-experience/
Source: Knowledgespeak

Deep Machine Learning Completes Information about the Bioactivity of One Million Molecules
A newly developed tool predicts the biological activity of chemical compounds, key information to evaluate their therapeutic potential. Using artificial neural networks, scientists have inferred experimental data for a million compounds and have developed a package of programs to make estimates for any type of molecule.
https://www.sciencedaily.com/releases/2021/06/210628132154.htm
Source: ScienceDaily

SciBite's new AI-driven Semantic Search Platform to help manage the surge in Life Science Data
Semantic technology company SciBite has announced the launch of its next-generation scientific search and analytics platform, SciBiteSearch. The platform offers powerful interrogation and analysis capabilities across unstructured and structured data, from proprietary as well as public sources. Researchers are currently facing challenges around accessing and deriving meaningful insights from the ever-larger volumes of data, presented in an array of formats from multiple sources. SciBiteSearch offers scientists access to domain specific ontology and AI-powered search capabilities, allowing users to connect and build knowledge from their data.
https://www.knowledgespeak.com/news/scibites-new-ai-driven-semantic-search-platform-to-help-manage-the-surge-in-life-science-data/
Source: Knowledgespeak

Pistoia Alliance launches DataFAIRy to drive AI Adoption
The Pistoia Alliance has announced the launch of the second phase of its DataFAIRy: Bioassay project, which aims to convert bioassay data into machine-readable formats that adhere to the FAIR guiding principles of Findable, Accessible, Interoperable and Reusable. The current pilot phase has been sponsored by AstraZeneca, Bristol Myers Squib, Novartis and Roche, and has successfully annotated 496 assays using a Natural Language Processing model that has been custom-built to recognise life sciences language. This second phase aims to scale the annotation process by 10 to 100-fold, and eventually promote the data model to become the industry standard.
https://www.scientific-computing.com/news/pistoia-alliance-launches-datafairy-drive-ai-adoption
Source: Scientific Computing World

eLife partners with PREreview to increase Diversity in Scholarly Peer Review
eLife, a non-profit organisation created by funders and led by researchers, has announced a new partnership with PREreview to engage more researchers from diverse backgrounds in peer review.
https://www.knowledgespeak.com/news/elife-partners-with-prereview-to-increase-diversity-in-scholarly-peer-review/
Source: Knowledgespeak

Optibrium and BioPharmics Partnership on Computational Drug Design
Optibrium and BioPharmics, providers of software and computational solutions for drug discovery, have announced a partnership to accelerate ligand-based drug design. The collaboration will combine industry-leading computational design solutions with advanced interactive user interfaces, engaging with experimental and computational drug discovery experts alike. The first outcome will be a new module for StarDrop, Optibrium's small molecule design, optimisation and analysis software, used by more than 160 organisations worldwide, including five of the top ten global pharmaceutical companies. The new Surflex eSim 3D module, planned for the 3rd quarter of 2021, will leverage BioPharmics' three-dimensional (3D) ligand-based design approaches, offering exceptional virtual screening and compound design capabilities for targets without available 3D structure information.

https://www.scientific-computing.com/news/optibrium-and-biopharmics-partnership-computational-drug-design
Source: Scientific Computing World

Elsevier collaborates with LexisNexis to increase access to Patent Information in Pharma and Chemical R&D Workflows
Elsevier has announced a new collaboration with LexisNexis Legal & Professional to strengthen the existing patent coverage in its information solution for chemistry R&D, Reaxys®. This integration gives companies and researchers access to the LexisNexis Intellectual Property patent content that powers LexisNexis PatentSight (Patent Analytics) and LexisNexis TotalPatent One (Patent Search) in their existing workflow.
https://www.knowledgespeak.com/news/elsevier-collaborates-with-lexisnexis-to-increase-access-to-patent-information-in-pharma-and-chemical-rd-workflows/
Source: Knowledgespeak

AI-based Clinical Trials Solution Provider Market Size to reach $5.2 Billion by 2028
Market is expected to expand at a CAGR of 21.7% over the forecast period, according to a new report by Grand View Research, Inc. Rising adoption of AI-based platforms and technologies by pharmaceutical companies and academia, along with increasing adoption of artificial intelligence (AI) are among the factors driving the market growth.
https://www.prnewswire.co.uk/news-releases/ai-based-clinical-trials-solution-provider-market-size-to-reach-5-2-billion-by-2028-grand-view-research-inc--824407866.html
Source: Cision PR Newswire

Springer Nature acquires Atlantis Press
Springer Nature has acquired Atlantis Press, an open access publisher specialising in conference proceedings and journal publishing, founded in 2006 in Paris. Containing oevr 1,200 conference proceedings, the Atlantis Press portfolio is seen to be a valuable addition to Springer Nature's book portfolio. It will further strengthen Springer Nature's position as a leader in the conference proceedings market, where it already publishes over 1,800 titles a year, and its leadership position in open access publishing.
https://www.knowledgespeak.com/news/springer-nature-acquires-atlantis-press/
Source: Knowledgespeak

Repurposing Drugs to Manage Covid-19 Progression
Researchers from the University of Cambridge, European Molecular Biology Laboratory's European Bioinformatics Institute (EMBL-EBI), VA Boston Healthcare System and Istituto Italiano di Tecnologia (IIT) have used a large-scale human genetics study to identify drug targets important for managing Covid-19 in its early stages. The research was able to identify the IFNAR2 and ACE2 proteins as targets for potential drugs for early management of Covid-19. While previous studies had focused mainly on identifying treatments for patients already hospitalised or critically ill, this study aimed to identify drugs that could be repurposed for early management of Covid-19.
https://www.scientific-computing.com/news/repurposing-drugs-manage-covid-19-progression
Source: Scientific Computing World

AAAS updates Author Name Change Policy
The American Association for the Advancement of Science (AAAS) has updated its author name change policy to support the privacy of authors who decide to change their name on research previously published across the Science Journals. The new policy, which went into effect February 18, 2021, is designed to support the privacy of authors. Authors may have occasion to change their names for multiple reasons. Recently, outreach by and on behalf of transgender scientists has impressed upon AAAS the importance of respecting authors' privacy and autonomy in correcting the scientific record. With this move, AAAS seeks to provide an inclusive research environment, which must be fostered in order for the publication and dissemination of excellent scientific research to continue.
https://www.knowledgespeak.com/news/aaas-updates-author-name-change-policy/

Source: Knowledgespeak

Springer Nature and the University of California join together to better understand Author attitudes to Open Research

New pilot to examine author awareness of open research principles and their benefits, and potential barriers to implementation.

https://group.springernature.com/gp/group/media/press-releases/springer-nature-uc-pilot-to-understand-open-research-attitudes/19067212

Source: Springer Nature

Machine-generated content: Boon to Information Professionals or the end of the world as we know it?

Advances in machine-generated content are changing scholarly publishing and Springer Nature is in the forefront of adopting its usage. What does this mean for researchers, scholarly publishing and librarians?

https://www.infotoday.eu/Articles/News/Featured-News/Machine-generated-content-Boon-to-Information-Professionals-or-the-end-of-the-world-as-we-know-it-144522.aspx

Source: Information Today

Yewno Adds News Content to Yewno Discover AI Research Platform

Yewno, a leading provider of Artificial Intelligence solutions, announced the enhancement of the Yewno Discover research platform to include news content.

https://www.infotoday.eu/Articles/News/Featured-News/Yewno-Adds-News-Content-to-Yewno-Discover-AI-Research-Platform-144481.aspx

Source: Information Today

US judge rules against GSK in Trade Secrets Dispute

Firm will not receive compensation for claimed loses of over $1bn from theft of confidential information by two former scientists.

https://www.chemistryworld.com/news/us-judge-rules-against-gsk-in-trade-secrets-dispute/4013897.article

Source: Chemistry World

Fraunhofer INT and Digital Science Collaborate

The Fraunhofer Institute for Technological Trend Analysis (INT) and Digital Science are collaborating to enable the technology foresight tool KATI (Knowledge Analytics for Technology & Innovation) to be combined with Dimensions data covering 116 million publications and 1.4 billion citations. They are also working together to make KATI available for users and clients outside of the Fraunhofer Gesellschaft. The newly formed collaboration was launched at the Hannover Messe 2021.

https://www.digital-science.com/press-release/fraunhofer-int-and-digital-science-collaborate-to-bring-fraunhofers-technology-foresight-tool-kati-together-with-dimensions-data/

Source: Digital Science

Exscientia acquires Personalised Medicine AI pioneer Allcyte

Exscientia, a clinical stage pharmatech company, has entered into a binding agreement to acquire Allcyte, a leader in artificial intelligence (AI) based precision medicine. The combination expands Exscientia's translational capabilities by enabling high content evaluation of individual patient biology in primary tumour tissues, rather than artificial cell lines or animal models. Allcyte's platform has been validated on multiple solid and haematological tumour types as well as in non-cancerous tissues.

https://www.exscientia.ai/news-insights/exscientia-acquires-personalised-medicine-ai-pioneer

Source: Exscientia News and Media

The most Discussed and Shared Research and Commentary of 2020

Covid-19, racial justice, climate change, the origins of life and other crucial issues all garnered attention in 2020.

https://www.digital-science.com/press-release/the-most-discussed-and-shared-research-and-commentary-of-2020/
Source: Digital Science

£250 Million Additional Funding to boost Collaboration and protect Ongoing Research
Scientists and researchers will get an extra £250 million funding this year to support pioneering research and drive the UK's ambitions to become a science superpower.
https://www.gov.uk/government/news/250-million-additional-funding-to-boost-collaboration-and-protect-ongoing-research
Source: Gov.uk

The life of a Chief Scientific Officer
You'll need a love of teamwork, an appetite for risk and an eye for strategy.
https://www.chemistryworld.com/careers/the-life-of-a-chief-scientific-officer/4013273.article
Source: Chemistry World

State-of-the-art Robot seeks out Chemical Agents
Dstl scientists have successfully trialled a fully autonomous robot that will help defence and security personnel dealing with hazardous scenes.
https://www.gov.uk/government/news/state-of-the-art-robot-seeks-out-chemical-agents
Source: Gov.uk

Solving a Puzzle to Design Larger Proteins
A team from Japan and the United States has identified the design principles for creating large "ideal" proteins from scratch, paving the way for the design of proteins with new biochemical functions.
https://www.sciencedaily.com/releases/2021/06/210624114416.htm
Source: ScienceDaily

X-rays Activate Cancer Drug
Radiation removes molecular mask to free anti-cancer agent, offering simultaneous chemotherapy and radiotherapy with fewer side effects in mice.
https://cen.acs.org/biological-chemistry/cancer/X-rays-activate-cancer-drug/99/web/2021/06
Source: C&EN

Chem Industry Overhauls Signal Changes for R&D
The COVID-19 pandemic has had a profound impact on nearly every industry, but Deloitte's new 2021 Chemical Industry Outlook suggests the chemical industry will see monumental changes as a result.
http://www.copyright.com/blog/chem-industry-overhauls-signal-changes-for-rd/
Source: CCC