Chemical Information & Computer Applications Group Chair's Report

Contributed by RSC CICAG Chair Dr Helen Cooke, email: helen.cooke100@gmail.com

2016 has seen the CICAG Committee working on a diversity of activities to support our members, including scientific meetings, sponsoring a student bursary and the Tony Kent Strix Award, building our social media presence. You'll find articles about all of these within this Newsletter.

During the second half of the year, we organised two successful scientific meetings in Cambridge: Chemistry on Mobile Devices and Cheminformatics for Drug Design (the latter in collaboration with the Society of Chemical Industry). We are finding that partnering with other RSC interest groups, and external organisations (the joint meeting with the SCI being an example), can be very worthwhile, providing opportunities for developing shared interests, widening our network of potential speakers and broadening our audience for events. With this in mind, in the next couple of years we are planning events with the RSC's Molecular Spectroscopy and Historical interest groups, and also external organisations and programmes including Dial-a-Molecule and the Chemical Structure Association Trust. If you know of, or are a member of, other groups or organisations who might like to work with us on events or other activities concerned with chemical information and computer applications please get in touch – we are keen to explore new opportunities.

We are also seeking to broaden our outreach by holding some events outside London. In 2017 we will be running a meeting on 22 June in Liverpool, entitled Chemical Structure Representation: what would Dalton do now? More information about this is provided in a separate article in this Newsletter.

We are also aware that not everyone can travel to face-to-face meetings, so for 2017 we are also exploring the possibility of [illegible]

RSC CICAG Social Media

Contributed by Dr Michelle Lynch, RSC CICAG Social Media Editor, email mklchems@gmail.com

CICAG runs a number of social media feeds including LinkedIn, Twitter and MyRSC. In 2017 we will add a Facebook group to this list. The current social media platforms broadly serve two important purposes - one being to publicise information specifically related to RSC CICAG's activities , e.g. meetings, events, news, publications and announcements, and the second to disseminate web links on topics relevant to our members and the wider chemical information and computer application community.

Our social media feeds are open to individuals and organisations with interests and activities that fall under RSC CICAG's remit as described above. Only RSC CICAG MyRSC is restricted to RSC members - one does not need to be a member of the RSC or RSC CICAG to follow our Twitter feed or to become a member of the Linkedin group.

Our Twitter feed had its five year anniversary in December 2016. It has grown both in numbers and in its coverage of not only traditional information science and chemical information publications, but also more aspects of computational chemistry, cheminformatics, advanced and digital manufacturing.

Taking a look at our Twitter statistics - we now have 373 followers, up from 229 in December 2015. Key topics of interest for our followers are science and technology news, biology, chemistry, biotechnology and medicine. Almost half of our followers (48%) are based in the UK, with the second most popular location being the USA (16%) and in third place India (6%).

The popularity of our tweets is always of interest to us. As of writing, we've tweeted around 200 times in total since the beginning of September. Some popular tweets from the last 3 months are shown in the table below, sorted by engagements. Twitter engagements include the number of times the audience clicks on the pictures, links or hashtags in the tweet, the RSC CICAG profile, or quotes, replies to or retweets the tweet. Twitter impressions relate to the number of times the tweet has been delivered to a user's account. A quote or retweet from a user with more followers therefore results in higher total impressions.

Recent articles with high popularity as judged by the tweet's exposure statistics relate to stories about the RSC Time4Chem initiative. Time4Chem was intended to be an opportunity for RSC members to help celebrate the RSC's 175th Anniversary by dedicating 175 minutes of their time to a chemistry related activity and to share their story. Our top recent tweet is an interview by the RSC with chemist and active Twitter user, Dr Nadine Borduas who tweets on the subject of "RealTimeChem". RealTimeChem is quite literally the study of chemistry measured in real time, often on mobile devices. Perhaps not surprising that our audience and followers show a preference for such inspiring stories about "real people" in our community; live coverage with photographs of RSC CICAG's events also stimulates a high number of Twitter engagements.

The growing importance of artificial intelligence and advanced manufacturing is reflected in our Twitter statistics, while traditional topics of big data, drug databases and libraries, chemical publishing trends and information for chemical education all still feature heavily.

We hope our Twitter followers and other social media members have enjoyed a positive experience of RSC CICAG social media in 2016 and would love to hear from you with any suggestions for topics to include for 2017.

Top 20 Most Popular Tweets Sept- Dec 2016 on the @RSC_CICAG Twitter Feed

#Where #Time4Chem meets #RealTimeChem	957	45
The Future of Artificial Intelligence and Cybernetics #AI	455	32
RSC CICAG Conference: Chemistry on Mobile Devices off to a great start! #RealTimeChem	482	27
A hands-on approach to chemistry with 3D printed molecules #3Dprinting #Chemspider #chemed	986	22
All together now: Why research papers have so many authors	779	20
#Insilico study of in vitro #GPCR assays by #QSAR modeling	260	17
Batch and flow: united at last #flowchemistry #realtimechem	922	16
Keynote Speech by Les Ebdon Director from Office for Fair Access on #diversity in Chemical Education #RSCGA16 #chemed	461	16
‘Robo-chemist’ optimises reactions in one day #flowchemistry #realtimechem #robotics #catalysis	1205	15
Our journals celebrate the winners of the 2016 Nobel Prize in Chemistry	639	14
Lufthansa Technik Flies with Mindbreeze InSpire #bigdata #KM	475	14
From haystack to needle: finding value with #DNA encoded library technology at GSK	732	13
York University gifted £1m for ‘world-leading’ chemistry research #chemed	420	

see the structures of the molecules. In collaboration with Luc Patiny at http://www.cheminfo.org/ we have been able to provide a visualiser that pulls data directly from the spreadsheet.

In addition there is a Vortex script (http://www.macinchem.org/reviews/vortex/tut26/scripting_vortex26.php) that pulls the live data from the Google doc and populates a Vortex workspace. More recently we have created a Jupyter Notebook (http://www.macinchem.org/reviews/osm/osmipython.php) and DataWarrior (http://www.openmolecules.org/datawarrior/) has been scripted to allow automatic import of the Google doc spreadsheet.

Whilst the spreadsheet has been absolutely critical to providing access to the OSM data, it is apparent that the limitations of a spreadsheet are going to be an issue in the future and the data needs to be moved to a database. The OSM team is now looking for help building an open-source chemically intelligent database, ideally with a web interface to allow collaborators around the world access. So if you feel you can contribute feel free to get in touch via the OSM website (http://opensourcemalaria.github.io/NewSite/#).

Tony Kent Strix Award

Contributed by Dr Michelle Lynch, RSC CICAG Social Media Editor, email mklchems@gmail.com

The UK eInformation Group, Geological Society, Burlington House, London, 31st October 2016

Picture courtesy of CILIP

The Tony Kent Strix Award is presented at the Strix Annual Lecture which is organised by the UK eInformation Group (UKeiG) in conjunction with RSC CICAG, the British Computer Society Information Retrieval Specialist Group (BCS IRSG) and the UK Chapter of the International Society for Knowledge Organisation (ISKO UK), in recognition of outstanding contributions to the field of information retrieval.

Robertson's presentation charted the timeline of information retrieval and described the forces that shaped the modern search engine followed by a look at how search engines use data. Early indexing systems all employed some form of standardised process (e.g. hierarchical or faceted classification) with controlled vocabulary and human assignment indexing. Mechanisation during the 1950-1960's existed mostly in the form of punch cards using Boolean principles. Towards the 1970's, computers were brought into the printing process by the large scientific abstract suppliers. Document digitisation at these companies became a key enabler of computerised searching, making it possible to carry out Boolean searches of both indexes and unstructured text. These enhanced search features expanded to be a commercial service provided by the scientific abstract and accumulator companies such as Dialog. Initially, computerised Boolean searches were performed by dedicated libraries off-

Read the full article here: http://www.rsc.org/news-events/articles/2016/dec/john-dalton-anniversary-celebration/.

RSC General Assembly, Europa Hotel Belfast 18th November 2016

Contributed by Dr Michelle Lynch, RSC CICAG Social Media Editor, email mklchems@gmail.com

The RSC's General Assembly (GA) is the main annual event for representatives of the RSC membership and governance committees. It also incorporate(l)-2e12(t)d -

Professor Les Ebdon, Director from the Office for Fair Access (OFFA), gave a very interesting keynote speech covering the need for diversity in chemical education. OFFA’s remit is “to promote and safeguard fair access to higher education for people from lower income backgrounds and other under-represented groups”. OFFA’s activities cover all full and part time undergraduate courses as well as teacher training courses in England.

Professor Ebdon spoke about the current difficulties in getting a representative proportion of female students as well as students from poorer socio-economic groups to study chemistry. Strategies proposed for improving diversity include “tailored, evidenced based spending”, collaboration and a “student lifecycle” approach which engages the “whole institution”. The sharing of responsibility was put forward with educational bodies, government departments, professional institutes, employers, unions, charities, families and student support functions all taking a role in improving diversity.

The RSC has played a key role through its campaigning and outreach initiatives such as “Chemistry for All” (Chem4All). Chem4all is an initiative being undertaken with several UK universities, local secondary schools and the Institute of Education. It is a five-year project that runs between 2014-2019 following secondary level students through years 8-12 of their education, and aims to overcome barriers to chemical education, widen participation in chemistry, and raise pupil awareness of their career options

Discussion Sessions

Membership Survey

The RSC membership survey was carried out to understand members’ perceptions of the RSC’s performance as well as their satisfaction and engagement rates. The information has been fed back into the RSC’s strategy and communication processes. The results of the RSC membership survey were presented and discussed in this session. The surv2(l)-3(mi)-2(s)b10(l)-2(35)6.j0 Tc63 Tw 0.349 fr)-2(o m(e m)e5o-h7(d)-2(s)o)2(m)(2)-10(of)--2(d)-ce

CICAG's Dr Diana Leitch is presented with the RSC Award for Service (Left to Right: Sir John Holman, Dr Diana Leitch, Dr Robert Parker); © Royal Society of Chemistry/MPP

CICAG Planned and Proposed Future Meetings

The table overleaf provides a summary of CICAG's planned and proposed future scientific and educational meetings. For more information, please contact CICAG's Chair, Dr Helen Cooke (helen.cooke100@gmail.com).

Meeting	Date	Location	Further Information
Chemical Structure Representation: What Would Dalton Do Now?	22 June 2017	University of Liverpool	See the meeting web page Jointly sponsored by CICAG and the CSA Trust
RSC Historical Collection	2017	Webinars	In partnership with the RSC Library and Historical Chemistry Interest Group

- Those working with complex structural information
- Providers of chemical data and information management software
- Chemical information publishers

This meeting will be sponsored by the Chemical Structure Association Trust. To find out more, please visit the RSC's Events database - this page will be updated as further details are confirmed.

CIIPM Training Events in 2017

Contributed by Stuart Newbold, email: stuart@psandi.com

At the 1st Cambridge Information & Intellectual Property Meeting (CIIPM), which took place at Homerton College, Cambridge, on 13 July 2016, feedback encouraged us to not only repeat the summer meeting, but also showed strong support for IP and patent information related training events. To this effect we are planning our first four half -day training sessions. These will run on 22 and 23 March 2017. The four training events must be booked separately, with discounts available for multiple bookings. The training event topics are:

1 IP Matters

- Provides an introduction to the basics of IP and the types of protection available, how to build an IP portfolio, and how to search and retrieve the various forms of published IP.

2 IP and Commercialisation

- Introduces the business and management processes needed to successfully get your technology to market including partnering, licencing, and IP portfolio management.

3 Commercial Patent Search & Analysis

-

The session aims at shedding light on the activity of chemists invested with responsibilities in the IUPAC, whose actions are often overlooked in national biographical dictionaries. This session falls into the broader project on the centennial of the IUPAC, in 2019. Each paper will focus on the responsibilities and actions of individual chemists, alone or combined in a small national or disciplinary group, inside IUPAC. The case study can however expand on roles in other international organisations (IRC, ICSU, SDN or UNESCO to name but a few). A first survey is provided by the books of R. Fennell (1994) and S. S. Brown (2001). By focusing on individual actions, the aim is to get a better sense of articulation between the local and the international, and how this articulation was constructed through the work and actions of chemists dispersed across the world.

Abstracts and a short CV can be sent to Danielle Fauque (danielle.fauque@u-psud.fr) and Brigitte Van Tiggelen (vantiggelen@memosciences.be)

The Working Party (WP) on History of Chemistry (EUCHEMS)

The meeting will also mark the fortieth anniversary of the creation of the Working Party (WP) on History of Chemistry of the European Association for Chemical and Molecular Sciences (EuCheMS). The WP business meeting will take place during the conference, and social events such as excursions, receptions, and a banquet will also enable celebration of the 40th anniversary to take place.

Specific presentations concerning the history of chemistry include:

- Hasok Chang (University of Cambridge): "What history tells us about the nature of chemistry"
- Maria Rentetzi (National Technical University of Athens): “Revising the concept of safety culture in nuclear settings”
- Anders Lundgren (Uppsala Universitet): “Science in chemical industry – what did it do?”

Noordwijkerhout Meeting 2017 Postponed

Due to logistical issues, the board of the Noordwijkerhout Foundation have made a decision that the initially-planned 2017 International Conference on Chemical Structures will instead take place in 2018. More details will follow in CICAG’s Summer Newsletter.

Reaxys and ScienceDirect at the RSC Library

As of 1 January 2017, access to Reaxys and ScienceDirect at the RSC Library has, regrettably, been discontinued. The decision to dispense with Reaxys is attributed to both increasing costs of the service

InfoChem News

Contributions from Dr Stephanie North and Dr Valentina Eigner-Pitto

Professor Nick Greeves, (Department of Chemistry, University of Liverpool), gave an interesting talk on mobile devices in chemistry education i

The history behind the 'Handbook of Medicinal Chemistry and Medicinal Chemistry Toolkit app - tablet delivery by design was described by Dr Andy Davis (Projects Director, RIA Innovative Medicines, AstraZeneca). The project started as an extension to the very popular Medicinal Chemistry training course organized by the RSC, initially as an e-book with links to external resources. This then evolved into a stand-alone iPad app (https://itunes.apple.com/gb/app/medicinal -chemistry-toolkit/id910073742). The Medicinal Chemistry Toolkit app is a suite of resources to support the day to day work of a medicinal chemist.
Based on the experiences of medicinal chemistry experts, these otherwise difficult-to-access tools are now available in a portable format for use in meetings, on the move and in the lab.

It is important to note that the calculations are done locally on the iPad and not via a web service. This is important for confidentiality when dealing with proprietary compounds. Properties are calculated as you sketch in the molecule so you can actually see the contributions of functional groups as you add them.
The Cheng-Prusoff equation, which relates IC50 values to affinities, is probably the most important equation for medicinal chemists involved in the design of inhibitors to understand.

As an example, for kinase inhibitors the Ki will vary depending on the concentration of the substrate ATP. It is therefore vitally important to understand the likely range in ATP concentration in the target locations, and to be aware that accumulation of the substrate (ATP) will have an impact on inhibition. This can be part of the reason why shifts in activity are observed when moving from a biochemical assay to a cell-based assay. Andy then asked for feedback on possible future directions. People have asked for it to be ported to other platforms, the ability to add extra functionality, user defined filters, other ADME models., all of which are possible but it was not clear how this might be funded or supported in the future.

Dr Dan Ormsby (Principal Application Scientist, Dotmatics Limited) gave a talk originally entitled "Challenges of developing electronic lab notebooks for mobile devices and substructure searching on iPad" but which was updated to include new developments on connecting laboratory instruments and devices into an electronic lab notebook environment. Elemental (https://itunes.apple.com/app/elemental/id518655328) is a chemical drawing package that is available on multiple platforms and was designed to provide the same user experience on both desktop and mobile devices. Whilst the chemical intelligence is all written in C for portability, the interface uses javascript to provide a consistent user interface.

Elemental has been downloaded many 100,000's of times and accounts for 15% of the links to the company website, underlining the role of small mobile apps as advertising aids. As an extension ElementalDB (https://itunes.apple.com/us/app/elementaldb/id627422287) provides a demonstration of an iPad application that does a substructure search of a 1,500,000 structure database in less than a second. The sdf records have been compressed to reduce size and put into a sqlite database. They then use the classing of path-based fingerprints for the structure- based searching, and calculate some physicochemical properties.

Isabelle Giraud (Actelion) then described DataWarrior (http://www.openmolecules.org) a freely available

pharma companies and applied a matched molecular pairs analysis (MMPA) to define a set of rules for favourable transformations (http://dx.doi.org/10.1016/j.drudis.2013.03.003). They can apply these rules to novel targets to suggest new compounds that would be worth making in an unbiased manner. Their analysis showed that whilst the data from an individual company provided a very limited set of useful rules, combining data from multiple sources resulted in