Statement of Need
for a National
Chemical Database Facility

Submitted by:-

Professor Rod Hubbard (York)
Professor Paul Raithby (Bath)
Professor Ben Davis (Oxford)
Professor Jeremy Frey (Southampton)
Dr. Amparo Galindo (Imperial College London)
Dr. Andrew Fogg (Liverpool)
Dr. Tim Claridge (Oxford)

1. Please state the type of facility and give a *brief* description of its function.

A National Chemical Database Service is required to provide a single, networked point of access to up-to-date, high quality, comprehensive chemical databases and ancillary facilities. It should also provide help, support, training and advice, trial new database systems and act as a proving ground for new mechanisms of data access.

This will ensure that the growing body of information derived from and essential for chemical research is conveniently accessible to a broad range of the EPSRC community. It will provide the important datasets and research informatics which are crucial to the understanding, enhancement and production of new materials and processes.

Its main target should be the UK academic research community, but its facilities should also be available for use in research training at all levels, ensuring that future as well as existing research communities are equipped to exploit such resources.

"The crystallographic databases are absolutely vital and heavily used by my group. The thermophysical properties database likewise." (Professor Chris Hall FRSC, Professor of Materials, University of Edinburgh)

"A central Chemical Database facility is essential to support a wide range of research and training activities in chemical processing, chemical engineering, chemistry etc." (Dr Megan Jobson, University of Manchester)

"A central chemical database is absolutely essential...There is no way that the UK can maintain its excellence in chemical research without such a service." (Dr John Fielden, University of Manchester)

2. Is this an existing UK facility or is it a new facility? If it is a new facility, please explain why this facility is now needed or will be needed in the future.

The Chemical Database Service (CDS) has operated comprehensive database facilities over the past 25 years. Continuing healthy growth in usage demonstrates such facilities are still considered essential by the UK academic community.

It currently includes Structural, Physical Property, Spectroscopy and Organic Chemistry Databases. A brief summary of these is provided under question 7.

3. What facilities of this type already exist (a) at the university level, (b) at the national or regional level and (c) at the international level. How accessible are these existing facilities to UK academics?

(a) At the University level

Currently, as far as we are aware, only the Cambridge Crystallographic Database is licensed for local workstation use by some users. However, such users still access the CDS as they can be sure the database is up-to-date and also contains the latest updates from e-Crystals (an archive for Crystal Structures generated by the Southampton Chemical Crystallography Group and the EPSRC UK National Crystallography Service).
None of the other crystallography databases or any of the spectroscopy or physical chemical databases are currently available at University level.

(b) At the National or Regional level

None of the databases listed in Section 7 are available at national or regional level (apart from via the CDS).

(c) At the International level

Individual accounts and, in some cases, Institutional accounts are available on international servers for individual databases such as the ICSD database, the Detherm database and the spectroscopy databases. Some of the front ends are as supplied to the CDS by the data producers, but the CrystalWeb interface is unique to the CDS, having been written in-house and allows simultaneous access to all crystallography databases. Access to many international services is based on IP address, which may bar work from home and, in addition, the usual 'pay by search' model will inhibit most users from using these services (see Question 5). None of the suppliers offer local training or offer Institutional ID (Shibboleth) logins to their databases.

4. Please describe who will benefit from the existence of this facility, including the number and type of researchers in the UK who are likely to want to use it and the research disciplines that it will benefit. Please indicate what level of usage such a facility would get in a year.

The whole of the EPSRC community can benefit from the existence of this facility.

For instance, the CDS currently provides full service support to:-

- Over 4,500 registered users from the UK academic community.
- 98 Universities, Research Councils and Institutes
- 410 different departments.

These departments range from chemistry (organic, inorganic, physical, biological, materials, medicinal, solid state etc,) to engineering (chemical, electrical, mechanical etc.), through biochemistry, biology, earth science, geology, pharmacy, physics etc. (see Research Interests in the Appendix)

In the last 12 months, on average there were:-

- Over 600 active users per month
- Over 4,400 accesses per month

Once logged in, users make several calculations per session on average. Many people/groups use the databases on a daily basis.

Usage of the CDS correlates well with the grade allotted by the Research Assessment Exercise. For example, users from chemistry departments rated in the 2008 RAE exercise (score >2) or the 2001 RAE exercise (Grade 3 or higher) accounted for:-

- 80.0% of all registered users
- 89% of all active users
- 83% of all accesses

The largest number of users and accesses came from Institutes with chemistry graded departments having scores of >=3 [followed by 2.95-2.75 then 2.7-2.5 then 2.45-2 then non-graded departments (or 5* > 5 > 4 > 3 in the old grading system)]

In the latest 2008 RAE, the top 25% of ranked Institutes (19 out of 41 in total - the others do not have chemistry or science departments) produce over half of the active users and accesses. This shows the high quality of research that the databases are supporting.

Since 1997, users have cited a marker CDS article (in J. Chem. Inf. Comput. Sci.) 854 times, which makes it the most cited paper in that journal (now J. Chem. Inf. Model.). A recent questionnaire by the CDS showed that users published at least 3 times more papers that do not cite or acknowledge CDS, but have still made significant use of the CDS. Most of the citing publications are in high impact factor journals and the wide variety of articles shows the impressive range of research disciplines that make use of the service.

The experience of the past five years suggests demand and use of such a central service will continue to grow, as all the scientific disciplines discover new ways to mine and exploit the constantly increasing

quantities of data and information. In addition, it is expected there is potential additional growth in users for new and extended services. For example, access was arranged recently for SPRESI, the world's third largest molecule and reaction database, at an extremely low cost via a central provision model. In the first few months, over 150 users have accessed this database.

5. Please explain why this facility is a "mid range facility" and what the benefits are of EPSRC supporting this facility. That is, why the facility needs to be supported at a national or regional level, rather than at a University or international level.

This facility can be defined as a mid range facility because:-

- The cost of the databases, software and hardware is relatively high.
- Only a central facility can combine and link data from different sources.
- Dedicated hardware at each University is not required.
- Expertise to mount the data, maintain it, make it available, provide front ends to the data and provide dedicated help, support and training may not be available at many (if any) University sites.
- Networked access to these databases would not otherwise be available at University, national or regional level.

"no one institution has the resources or ability to hold such a vast repository of information." (Dr Mark Russell, IT and ELearning Support Officer, Chemistry Department, Bath)

A central facility provides a 'one-stop-shop' for the data, obviating the need to access multiple sites. For instance, the CDS has produced a web front end (CrystalWeb) for *all* the crystallography databases, which is still undergoing active improvement and development. For instance, it may be extended to include simultaneous searching of non crystallography databases. This is unlikely to have happened at either an individual, university or international level.

A central facility is very well placed to trial new databases, as demonstrated when the CDS introduced Beilstein and Gmelin to the academic community, negotiating and running a 6 month long evaluation trial.

Help, support, training and advice can be offered much more efficiently by a central service with a dedicated team based in the UK rather than relying on individuals at Institutes, and it is certainly much cheaper than anything available from database suppliers. Almost all of the suppliers are based outside the UK and training courses are not available, nor do they have outreach programmes and rely on conferences and user meetings to interact with their users. The CDS is able to supply a lot of dedicated help and advice by e-mail, over the phone and on its web site. This includes tutorials, exercises and Flash movies that give details on how to access and use the databases.

The cost of a single national site licence is considerably less than the sum of individual site licences. For example, the current cost for a University site licence giving unrestricted ACD/LABS I-Lab for 12 months is £4,950. Over the last 12 months, users from 53 Universities and other Research Councils/Institutes have used this database. If all Institutes licensed it individually it would amount to a total cost of five times what the CDS has negotiated for a single national licence (£50k/year), or **81%** of the total cost that the EPSRC pays for provision of the whole of the database facility (see Question 7 – costs). Similar benefits of cost occur with all the other databases.

The CDS obtained a 'site' licence for the SPRESI database, the world's third largest molecule and reaction database for the same cost that an individual Institute would pay. During a 3 month trial, 45 Universities and other Research Councils/Institutes accessed SPRESI through the CDS. If they bought individual site licences then this would have amounted to 112,500 Euros. Although there are notional restrictions to the number of simultaneous users, this method is obviously extremely cost effective.

This also demonstrates the other advantages of a service like the CDS – that of having a single central point of access with expert help available.

For most of the databases, personal workstation copies are not available and must be run from dedicated servers. To run the databases at each University (if the suppliers permit it), the cost would be very high as several server machines (Windows and UNIX) would be required with the appropriate software installed (such as Oracle) which may require a separate licence. Also an experienced person (1/2 FTE? at each Institute) would be required to install and update the data as well as provide and maintain connectivity and tools such as web front ends for most of the databases as well as help and training. The cost and extra work involved would almost certainly mean that most Universities would not be able to do this. International service providers would be unlikely to be in a position to offer Institutional (Shibboleth) logins where as a national service can (and the CDS currently does).

"This must surely be less expensive and more efficient than any attempt to run databases simultaneously or independently from separate locations." (Professor A. L. Greer, Head of the Dept. of Materials Science & Metallurgy, University of Cambridge)

6. How long should the facility be supported for?

It will be required as long as there are research communities which require access to, and benefit from, chemical information. Effective access is essential to keep UK science at the forefront of international and world research and to be internationally competitive.

For example, the experimental information on thermodynamic properties found in Detherm is key in chemical engineering, promoting the transfer of theoretical developments into tools for the design and synthesis of chemical processes and products.

*"The CDS is for myself, my group and most people in my department an absolutely pivotal service. The EPSRC support for CDS should be continued, as it serves to support invaluable services to UK science. Without the support of this service we are more likely to slip behind other countries" (*Prof. Martin Wills, Professor of Organic Chemistry, University of Warwick)

In 1992 a committee was set up to review the whole issue of chemical database provision for the UK academic community. It concluded that there were a number of distinct benefits offered by a dedicated central facility. These benefits include cost efficiency, ready availability of specialist information all in one place, access to large systems beyond the means of individuals, user support and training.
This is still true today and will remain so in the foreseeable future.

The dream of Grid technology, where all useful chemical data is supplied to researchers from distributed open-access repositories, has not yet materialised. The reality is that a central facility like the CDS is needed now more than ever before and will be required for the forseeable future.

With the amount of chemical information growing ever larger, a central source to access this data becomes even more vital. This is shown by the fact that the numbers of users accessing the crystallography databases from February to May 2009 were the highest for any months over the last 5 years. Similarly, user numbers and accesses to the spectroscopy databases have increased 3 to 4 fold over the last year and are still rising. This facility is needed now and in the future to underpin and facilitate research and maintain UK academic research at the forefront of international and world research.

In the future, a central facility may be the best option in exploiting the various emerging institutional repositories by harvesting metadata and providing a common central source of access. The CDS already does this for data published by the eCrystals project.

7. Please indicate what the facility should provide to be of maximum benefit to the research community and estimate the likely cost of the facility. For example, indicate what size should it be, what technologies should it have available, how many staff would it need. You should prioritise these requirements in terms of "must have" and "desirable". In addition, please highlight any features that would be detrimental.

Must Have
The facility should provide unlimited access, via web or client software, to all of the databases that are currently made available (see section 2) as a must have. Server machines should be secure and well maintained and the facility should have excellent network connectivity and provide tools for interrogating, visualising and exporting the data. The data should be updated regularly in a reliable way.

The databases are

- The Cambridge Structural Database (organic and organo-metallic compounds.)
- The Inorganic Crystal Structure Database (inorganic compounds)
- The CrystMet database system (metals, alloys and intermetallics)
- The NIST Crystal Data Identification File (crystal class and unit cell dimensions.)

Physical Property Databases

- DETHERM (thermophysical property databases of pure compounds and compound mixtures. Contains nearly 6 million data sets for ~134,000 systems (~ 28,000 pure substances and ~106,000 mixtures) covering around 500 property fields.)
- ACD/Labs I-Lab (pKa, LogP and Solubility searching and prediction plus prediction of LogD, Boiling point/Vapour pressure/Enthalpy of vaporization prediction and Adsorption coefficient/ Bioconcentration factor plus naming of compounds)

Spectroscopy Databases

- The ACD/Labs I-Lab (^{1}H, ^{13}C, ^{15}N, ^{19}F, ^{31}P NMR spectroscopic searching and prediction)
- The SpecInfo spectroscopic system (^{1}H, ^{11}B, ^{13}C, ^{15}N, ^{17}O, ^{19}F, ^{31}P NMR spectroscopic searching and prediction plus IR Spectra and Mass Spectra)

Organic Chemistry Databases

- In-house small chemical procurement catalogue and screening compound databases.

Plus for a 1 year 'trial'

- SPRESI database, the world's 3rd largest chemical structures and reaction database.

Currently, the CDS provides added value by producing an in-house web interface, CrystalWeb that enables *simultaneous* searching and hits display for *all* the crystallography databases. It also provides links to the primary electronic literature as well as integration with the SpecInfo spectroscopy system.

There should also be, as a must and bare minimum, two staff (posts 1 and 2) with appropriate chemical, technical and computing expertise in order to run a quality service.

Professor George Jackson FRSC, Imperial College London said;-*"As chairman of the statistical mechanics and thermodynamics group (SMTG) and secretary of Faraday Council of the Royal Society of Chemistry, I would like to expression my strongest support for a national facility and the DETHERM thermophysical property database in particular. Thermophysical properties data of this type is indispensable not only to experimentalists in the physical and biological sciences, but also to theoreticians and molecular modellers. Physical chemists and chemical engineering require access to such data in their research programmes as a matter of routine… Research in this area would simply not be able to proceed in an effective manner without the ease of access to data provided by database such as DETHERM."*

"I am a crystallographer, primarily using crystallographic tools which are INVALUABLE through CDS". (Prof A M Z Slawin, Professor of Chemical Crystallography, University of St Andrews)

"A central chemical database is of enormous benefit to our research. We mainly use the Cambridge structural database and SpecSurf. I would regard these as essential." (Dr Alan Simm, School of Engineering, University of Edinburgh)

"*ACD/Ilab software, NMR prediction software, IUPAC naming software must have"* (Prof. Ron Grigg FRS, Professor of Medicinal Chemistry, University of Leeds)

Desirable

In summary:

- It would be **highly desirable** to have an additional post (post 3) to develop a common portal to all data and to have sufficient capacity to trial new databases.
- It would be **desirable** for a further post (post 4) to develop and deploy research training materials and to have the flexibility to extend the range of databases available.

It would be highly desirable to have a single portal to access all the data from the facility databases plus that harvested externally from the various emerging institutional repositories by harvesting metadata and providing a common central source of access. The CDS already does this for data published by the eCrystals project. This project currently includes an archive of crystal structures generated by the Southampton Chemical Crystallography Group and the EPSRC UK National Crystallography Service. It would be highly desirable to have an additional person (post 3) to develop such a facility to further combine, integrate and link data from different sources, internally and externally.

In addition, it is highly desirable that there should be provision to allow the facility to trial any new database that may become available in the future in order to assess the potential benefits for the UK academic community.

"A searchable database of commercially available compounds and screening sets would be extremely useful" (Prof. Alan Fairlamb, University of Dundee)

"the potential for identification of future databases, essential to competitiveness, will be lost without the capacity provided by the CDS to broadcast and appraise them" (Dr. Chris Garner, Nottingham Trent University)

It would be desirable to have another person (post 4) available to develop a research training programme. He/she could be deployed around the UK so that all researchers (faculty staff, post-docs and students) are fully trained to use chemical information tools and access the enormous wealth of data which exists and is available to them. Provision of enhanced training materials could also increase usage, making important contributions to the quality of research undertaken in the UK.

It would also be desirable to extend the range of chemical databases offered to include more specialist databases, such as the Available Chemicals Directory (ACD). This particular database is of use to a large community and was heavily used when previously available via the CDS (over 200 users per month). A survey run in 2008 showed that half of those that responded reported that the loss of ACD had a major effect on their research. After a year of stopping this database *88*% of responses showed that no alternative had been found and most said any alternatives were worse and harder to access. A facility of this kind not only saves money but also time and effort otherwise wasted by having to search numerous web sites.

"The database that we used to use regularly prior to its removal was the ACD (Available Chemicals Directory) - alternative sources of this information are much less satisfactory and lack of access to this database is really frustrating. (Prof. Alan Spivey, Imperial College)

Costs

The cost to the EPSRC for the current Service (for the 3 year period April 2008 to March 2011) is in the region of £300K a year. This covers the cost of all the database/software licences and hardware needed to run them plus the cost of 2 staff including overheads and other expenses such as travel and training. Costs for a future service would inevitably differ as, amongst other things, database costs are dependant on exchange rates. Addition of the SPRESI database under the current licence agreement would only add a further £2.5K per year. Inclusion of 'desirable' databases would add something in the region of £50K a year. Addition of 2 staff for developing database systems and for training would add an additional cost in the region of £150k per year.

"The central provision of a chemical database system is a highly cost-efficient way of making data available to the UK academic community" (Professor Sandy Blake, Professor and Director of Chemical Crystallography, University of Nottingham)

8. If EPSRC was unable to support this facility, what would the research community do? (for example, in terms of looking for other sources of financial support or seeking access to non-UK facilities)

The research community can be considered in three categories and the effect on them would be:-

1) **Expert** (daily-weekly user).
As access to the data is vital, this user would probably mount the data locally (only available for Cambridge Structural Database and Crystmet) or subscribe on an annual (or Institutional) basis to an on-line international service (if available), provided funding could be obtained (see below). This would prove to be very expensive and time consuming (setting up and maintaining the data) which would be to the detriment of the users research work.

2) **Frequent** (weekly-monthly user)
Again the data may be vital so this user would probably subscribe (monthly/weekly rate or per access) to an on-line international service (if available). This would prove to be very expensive and because of the cost and difficulty of access, the user would cut the number of searches to a minimum - to the detriment of his research.

3) **Occasional** (monthly or less)
This may also include students requiring access for research training purposes. The information may be vital to the work but this user would probably not bother to subscribe (or be able to afford a subscription) and therefore his research or training would suffer.

The following are some additional comments on the impact of not having the central service:

- Institutional licences are not always available, or are expensive (see Question 5) and the question of who pays for a site licence is hard to answer as a very broad range of disciplines can be involved (individuals, departments individually or collectively, libraries etc.).
- The research community might attempt to bundle licence costs into their grant applications to the EPSRC. Given the current low responsive mode success rate many users would loose out and, were they to be successful, the aggregate cost would be extremely high.
- Access to a database is not always straightforward. For example, at present, the Cambridge Structural Database has to be licensed and installed separately on individual workstations – with network access only via the CDS. *No other on-line access for UK academics to this database is available*.
- Funding of this type of facility by sources other than the EPSRC would be very difficult. It would not be possible to obtain funding from Industry as licence agreements preclude anyone except UK academics and, if it was extended, then licence costs would rise enormously.

Therefore, if the EPSRC decided to discontinue support for this type of facility the research community would find it much more difficult to access the data and more difficult to fund access to the data. Many would feel inhibited by the cost of conducting multiple searches and be deterred from extensive access to the data, leading to an overall reduction of quantity and quality of research.

"Providing such a service on a national basis is far more effective (including probably more cost effective to the EPSRC!!) than expecting each lab to individually purchase licenses for crystal / molecular structure databases from their individual research project funding." (Dr. Ian MacLaren. Glasgow Physics)

"I worry that loosing the CDS could be catastrophic not only to research but also for teaching outside of the grade 5 club." (Dr John J. Morrison, Manchester)

Professor Peter Willett (Sheffield) commented

"I am not a synthetic, physical or structural chemist, but I have been associated with one of the world's leading chemoinformatics groups for three decades. Charging the costs of database access to research grants via FEC has two problems associated with it. First, it would mean that only those who were successful in obtaining EPSRC research awards would be able to gain access to the data, and I would assume that the user base is far larger than the number of holders of current EPSRC grants. Second, it would mean that each successful applicant would have to pay for a full academic licence and to develop the expertise necessary to access the particular database(s) of interest. Even for a dedicated chemoinformatics group such as the one I work in, the plethora of data formats and command types mean that we restrict the number of systems with which we work on grounds of feasibility: this data variability is likely to be a significant problem for non-specialist users.

The service is not competing against commercial database providers. If this was correct then the providers would not be willing to make their data available for repackaging via CDS and subsequent processing - most conventional database licenses strictly forbid any such manipulations.

*Second, the impact of CDS is evident in a recent citation analysis of the most important chemoinformatics journal, the American Chemical Society's Journal of Chemical Information and Computer Sciences, which showed that the standard CDS reference (Fletcher, D.A. et al., J. Chem. Inf. Comput. Sci. 1996, 36, 746-749) has been the second most-cited of all the papers published in the journal in 2005 (see http://www.warr.com/25years.html).**

9. Please make any other comments that you think are relevant to the statement of need for the facility.

The EPSRC recently (2007/2008) put out a tender for these set of databases (Ref: R5-E3-34). This demonstrated that:

1) An EPSRC advisory panel already recognises that a number of databases are key requirements for the UK academic community.
2) The best way of supplying these databases is via a central facility.

Even more recently (February 2009), the importance of one of the databases, ACD/Labs, was confirmed when EPSRC decided to continue funding it for a further 2 years after providing funding for an initial one year assessment period.

Further confirmation of the importance and need for this type of facility is provided by the results of a user community canvassing exercise which was set up by the SoN sponsors to allow this community to express its views on the continued availability of facilities of the kind currently provided by the CDS.

Over 800 have expressed their support for such a facility, of which at least two thirds are postdoctoral researchers (staff or research fellows, almost 100 are Professors) from 85 Universities /Institutes/ Research Councils and almost 40% have supplied comments. Most of the comments say that a facility like this is either essential (78), vital (34), critical (8), very important (27), extremely or very useful (45) or [in]valuable (39) to their work.

Information on the above exercise can be found at:-

http://www.ipetitions.com/petition/ChemicalDatabaseService/index.html

The full list of names, Institutes, research interests and comments can be found in the Appendix.

This Statement of Need is obviously influenced by the existing EPSRC-funded Chemical Database Service and many of the statements of support directly reference the existing resource. It should be emphasised that what we are making the case for is a facility of this type.

10. Who was involved in preparing this statement of need? Please list name, institution and research interests.

This statement has been produced initially by the Management Advisory Panel to the CDS who broadly cover the research community that use the facility.

Members of this panel are:-

Professor Roderick E Hubbard, University of York. Chairman of the BioMolecular Sciences committee of the BBSRC.
Research interests have focused on understanding the relationship between structure, mechanism and function in various protein systems (including proteases, nuclear receptors and kinases) and experimental and theoretical studies of protein-ligand interactions.

Professor Paul Raithby, Chair of Inorganic Chemistry at the University of Bath. President of the British Crystallographic Association and Chair of the working party that is developing the new "small molecule" single-crystal X-ray diffraction beamline.
Research interests are in the area of solid state materials and materials science using a range of synthetic, spectroscopic, diffraction and knowledge mining techniques to probe the correlation between materials and their properties

Professor B.G. Davis, Department of Chemistry, University of Oxford.
Research interests include synthesis using both biotransformations and conventional synthetic methods to molecular modelling, enzyme kinetics and protein chemistry for the manipulation and redesign of naturally occurring carbohydrate-containing structures.

Professor Jeremy G. Frey, Professor of Physical Chemistry, University of Southampton. Research interests include Interfacial Spectroscopy, Nanoscale Laser X-Ray Scattering and application of modern computer science, semantic web, web 2.0 to chemical research.

Dr. Amparo Galindo, Reader in Physical Chemistry. Department of Chemical Engineering, Imperial College London.
Research interests are the development of statistical mechanical approaches for complex systems and their application to processes relevant to industry**.**

Dr. Andrew Fogg, Department of Chemistry, University of Liverpool.
Research interests are the synthesis and characterisation of new inorganic and inorganic/organic composite materials.

Dr Tim Claridge, Director of NMR Spectroscopy for Organic Chemistry, University of Oxford. Chair of the RSC NMR Discussion Group.
Research interests revolve around the application of solution-state NMR techniques to address questions of structure, function and dynamics of "small" molecules in organic chemistry and chemical biology.

The following have contributed to the text of this document:-

Professor Peter Willett, Professor of Information Science, Sheffield University.
Research interests: Bibliometric methods for the evaluation of research productivity; chemoinformatics, in particular the use of clustering, graph theory, and machine learning methods for the processing of databases of chemical and biological structures.

Professor George Jackson FRSC, Professor of Chemical Physics, Department of Chemical Engineering, Centre for Process Systems Engineering, Imperial College London. Chairman of the Statistical Mechanics and Thermodynamics Group (SMTG) and secretary of Faraday Council of the Royal Society of Chemistry.
Research interests: Statistical mechanical theories and computer simulation, phase equilibria of systems which are of industrial relevance, liquid crystal modelling.

Professor Chris Hall FRSC, Professor of Materials, The University of Edinburgh.
Research interests: chemistry of materials and geomaterials used in oilfield and civil engineering.

Dr Megan Jobson, Senior Lecturer, School of Chemical Engineering and Analytical Science, University of Manchester.
Research Interests: methodologies for conceptual design and optimisation of heat-integrated distillation systems, ternary and multicomponent azeotropic distillation sequences, absorption and reactive absorption processes, and reactive distillation processes.

Dr John Fielden, Research Associate, Department of Chemistry, University of Manchester
Research Interests: Inorganic chemistry, dye chemistry, cluster chemistry, materials chemistry.

Dr Alan Simm, School of Engineering, University of Edinburgh.
Research Interests: Bio-engineering.

Dr. Ian MacLaren, Solid State Physics, Physics and Astronomy Department, University of Glasgow.
Research Interests: Physics & Materials Science.

Professor Sandy Blake, Professor and Director of Chemical Crystallography, University of Nottingham.
Research Interests: Structure determination; low temperature crystallography; high pressure crystallography; supramolecular structures.

Professor Alan Fairlamb, CBE, FRSE, FLS, FMedSci, Professor of Biochemistry, College of Life Sciences, University of Dundee.
Research Interests: Target validation and drug discovery for parasitic tropical diseases

Professor Alan Spivey, Professor of Synthetic Chemistry, Imperial College London.
Research Interests: Organic Synthesis.

Professor Alexandra MZ Slawin, Professor of Chemical Crystallography, University of St Andrews.
Research Interests: Chemical Crystallography, Structure solution and structure activity relationships. Automation of crystallography.

Dr Mark Russell, IT and ELearning Support Officer, Chemistry Department, University of Bath.
Research Interests: Computational Chemistry and eLearning.

Professor Martin Wills, Professor of Organic Chemistry, Warwick University.
Research interests: Selective and active new catalysts for the asymmetric catalysis of organic reactions

Professor Ron Grigg FRS, Professor of Medicinal Chemistry, School of Chemistry, University of Leeds.
Research Interests: Cascade Reactions, Heterocyclic Chemistry, Palladium Catalysed Reactions, 1, 3-Dipolar Cycloaddition Reactions, Asymmetric Synthesis, Medicinal Chemistry, Combinatorial Chemistry

Dr John J. Morrison, Organic Materials Innovation Center, School of Chemistry, University of Manchester.
Research Interests: Porous and Auxetic Organic Materials.

Professor A. L. Greer, on behalf of the Deptartment of Materials Science & Metallurgy, University of Cambridge.
Research Interests: materials science, metallic and chalcogenide materials.

The following have also commented on this statement by private communication. Most of those below have expressed a desire to see databases like the Available Chemicals Directory available through a central service.

Professor Peter M. Fischer, FRSC, on behalf of the Division of Medicinal Chemistry & Structural Biology of the School of Pharmacy, University of Nottingham.
Research Interests: Organic and medicinal chemistry, computational chemistry, structural biology, and bioscience.

Professor Timothy C Gallagher, Professor of Organic Chemistry and Head, School of Chemistry, University of Bristol.
Research Interests: Heterocyclic synthesis, carbohydrate chemistry and the chemistry and pharmacology of Nicotinic Receptors.

Professor Alan Armstrong, Department of Chemistry, Imperial College London.
Research Interests: Synthetic Organic Chemistry: Methodology, Asymmetric Synthesis, Total Synthesis of Natural Products.

Professor Kenneth Douglas, FRSC, Director of the Wolfson Centre for Rational Design of Molecular Diagnostics, School of Pharmacy and Pharmaceutical Sciences, University of Manchester.
Research Interests: medicinal chemistry.

Professor David R. Adams, Professor of Organic Chemistry, School of Engineering and Physical Sciences, Heriot-Watt University
Research interests: Synthetic organic, biological and medicinal chemistry.

Professor Duncan Bruce, Professor of Materials Chemistry, University of York and on behalf, as Chair, of the Royal Society of Chemistry Materials Chemistry Forum.
Research Interests: Liquid Crystals and Materials Chemistry.

Professor Philip Mountford, FRSC, Professor of Chemistry and Executive Deputy Head of the Inorganic Chemistry Laboratory, University of Oxford.
Research Interests: Organometallic chemistry and catalysis.

Professor Robert. J. Deeth, Associate Professor and Reader of Computational Chemistry**,** University of Warwick.
Research interests: Inorganic computational chemistry

The panel has also taken into account the hundreds of comments from the user community canvassing exercise. The full list of names, Institutes, research interests and comments can be found in the Appendix.