



Engineering and Physical Sciences
Research Council

Chemical Database Service

Draft - Review Report

6th – 7th June 06

Prepared by: Dr Carmine Ruggiero

Date: 20th June 2006

1 Forward

1.1 Introduction

EPSRC has developed a cross service review format to examine all of the Chemistry National Services on an annual basis. In order to minimise the burden on services the primary information supplied to the panel was the existing service annual reports. Prior to this annual evaluation, the reports have been supplied to EPSRC and the individual Management Advisory Panels. It was considered that the reports should be the primary information for an on-going evaluation of services. Recommendations had been made to the services in the previous reviews and the panel also examined how the annual reports had been developed in light of these.

The dual objective of the panel meeting was to review the operation and function of the current national services and also assess the UK requirement for the mass spectrometry and chemical database national services. The panel met over 2 days and prior to assessing the requirement for the two national services, there were light touch reviews of the EPR, Solid State NMR, X-ray crystallography and Computational Chemistry National Services.

The chemistry national services vary widely in terms of their history, the range and level of the services they provide, and the level and form of funding they receive. There has never been a specific initiative to establish such centres. The current services are a result of ad hoc processes, all of which were essentially demand led. Most of the service providers provide a 'full service', by which they are able to analyse samples, and produce and interpret the data. However, the amount of knowledge and time required to perform such a full service varies dramatically between services.

In convening this annual assessment, EPSRC is attempting to develop a more coherent strategy for assessing the services and this is the third such a cross service evaluation. The most recent cross service review took place in May 2005; the individual grades from this previous review are included for information. Annex 1 details the panel membership and terms of reference; annex 2 lists the services and service directors.

Detailed Objectives

1.2 Detailed Objectives

3 primary objectives:

- *Review each of the services based on existing renewal and annual reports.*

This light touch review of the services will be based on the literature supplied to EPSRC on an annual basis. The aim is to highlight potential issues with sufficient time to allow the service to address them prior to a future renewal. The panel are also requested to identify areas where the existing reports continue to lack detail. The services have received feedback from the previous meeting and are aware of this current review.

- *Assessment of the UK chemistry requirement of the chemical database and mass spectrometry chemistry national services.*

This panel membership reflects the breadth of chemistry supported by the chemistry programme including the main areas supported by the services. Prior to the meeting, the two services scheduled for a review were invited to provide a vision statement highlighting their rationale for a service and how such a service would complement the equipment and research landscape. The documentation was also sent out to two expert assessors, a single response was received for each service and their comment was also sent to the service for response.

The panel considered the case for the continuation of a chemistry national service against the published criteria for the service, in order to provide advice to the Chemistry Programme Manager on whether or not to proceed to announce an expression of interest to operate the service. The panel could also recommend specific elements that should be included in any call for proposals.

- *Provide overall health check and highlight best practice across all the chemistry national services.*

The panel also highlighted any issues that affected more than one of the services, including any areas of best practice that could be fed back to all the services.

1.3 Service Assessment Criteria

The individual reports contained within this report are structured around the assessment criteria. The panel comments for each service are collated under the following assessment criteria:

- Demand (primarily by academic chemists)
- Research Quality (of projects facilitated by the service)
- Efficiency and management of service
- Performance against service benchmarks (where applicable benchmarks will be included in the volume)
- Training (workshops / summer schools)
- Publicity (highlighting service to new users)

The panel were also asked to provide grades on service performance against each of the assessment criteria as well as an overall grade. The grades are based entirely on the information supplied to the panel and could simply reflect a communication issue which the service should address in subsequent reports. The traffic light grading system used by the panel is summarised below:

Green	Service performing well in this area, with no issues
Amber	Service generally performing well, but with a few minor issues or potential future issues
Red	Service not performing to standard, with major issues to be addressed in the coming year

1.4 Briefing documentation

Each service was assessed on the same briefing material.

- **Renewal report** – The panel report from the last service grant award. Each report should detail the main issues which the panel examined prior to funding the proposal. These reports may also list concerns which the panel recommended be addressed during the lifetime of the current grant.
- **Annual report** – Each service is required to supply an annual report which details how the service has operated during the past year including highlighted the usage and users of the service.
- **Vision document** – The chemical database and mass spectrometry services also submitted a vision statement detailed their rationale for the continuation of a national service for the next 5 years. The services included an assessment of how their service complimented the existing research and equipment landscape.

1.5 Next Steps / Actions

The report will be fed back to individual services, who will receive detailed comments on their own service and cross service recommendations. Actions from this review will also be discussed at Management Advisory Panel meetings. The aim is for future annual reports to provide sufficient information for subsequent panels to continue to provide an accurate service 'health check'.

A fourth cross service review meeting will be convened during 2007 to assess continued progress and outcomes from this initial review.

Actions for Services and EPSRC

- To ensure sufficient transparency of EPSRC funding, annual reports should be available via the service and EPSRC websites, the cross service panel report should also be made available.
- The EPSRC to continue to work with services in delivering publicity where appropriate. In addition to collate research highlights from all the national services and to develop for use in cross service publicity.
- Annual reports from each service to cover the period Apr – Mar, and should include MAP minutes and membership (including length of service on MAP).

Key elements for the services to consider and be included in future annual reports:

Demand

- Statistics for service usage against demand. The statistic should also be broken down to individual techniques where possible.
- An explanation as to how demand is managed and how demand relates to capacity
- Who is using the service and why is the service required (particularly for major users)
- A percentage breakdown of new users

Research Quality

- Research Highlights
- List of service outputs including presentations
- How critical was the service to a publication
- Have any service results lead to further research funding
- An indication of the breadth of research areas supported
- Provide data on individual institution / departmental usage
- Research supported at the interfaces of chemistry

Management

- MAP membership and minutes
- When MAP members were appointed and scheduled to step down
- How sample prioritisation has been carried out
- An indication of management structure within the service

Performance

- Each service to have benchmarks set and reviewed annually by the MAP
- Benchmarks should examine quality as well as quantity
- Benchmarks should initially detail minimum performance standards
- Service should provide data on turn around times and throughput of samples
- Instrument down time should be outlined

Training

- Number of applicants versus attendees for training elements should be detailed
- Feedback should be included
- Summary of training content
- Any informal training carried out and its demand

Publicity

- How are the services presenting to communities a range of potential user communities
- Additional information on what a particular technique can do should be available (application notes)
- Articles in Chemistry World or other appropriate publications

2 Chemical Database Service

2.1 Grant Summary

The National Chemical Database Service is accessible by all UK academics. Access is provided to a range of databases in the following four areas of chemistry. Crystallography, Synthetic Organic Chemistry, Spectroscopy and Physical Chemistry. Support and training, via courses and online material, is provided for users of the service. New technologies and products are acquired where appropriate to ensure maximum ease of use of the databases. A watching brief is kept for new sources of relevant data. The Service is publicised through a variety of media to ensure that all eligible users are aware of the potential of the service in supporting their research effort.

2.2 Initial comments

The panel discussion was more in-depth due to additional discussion regarding the case for the continuation of a service. The current service grant is 3 years in duration, and provides a portfolio of databases. The structural databases have historically had the most usage; DETHERM is the major new database purchased during the current grant and the service is assessing potential new databases on an ongoing level. The PI is supported by a four person team. The remit of the service is to support users across the chemistry community. Historically service publicity has primarily been through the website; this has now been complemented with an extensive series of institutional visits.

2.3 Demand

The structural databases are fundamental to the service and the demand for this is high; the demand for the rest of the portfolio is unclear. Examining the data there has been a 30% increase in users; however when users who are no longer active are taken away this is only a 5% increase. There has been a 35% increase in users from Chemical Engineering primarily these would be new users for DETHERM.

Following on from last year's cross service report, the level of CDS usage appears to be relatively static despite the ongoing publicity effort. The publicity programme has delivered some new users though not for the databases as hoped. The largest numbers of new users have been for the structural and DETHERM databases; the structural databases already had high usage, while DETHERM is a relatively new on-line database which is demonstrating good levels of access early in its lifetime. The synthetic organic and physical chemistry databases have had much lower increases, and these were the ones with relatively low usage where there was potential for significant increase.

Grade: Red

2005 Grade: Amber

2.4 Research Quality

The quality of research supported by the service continues to be difficult to assess. The research highlights appeared to be improved from previous years; however taking this as a sole benchmark of what the service is delivering is not sufficient. In examining the highlights three of the seven were based on structural data. The number of citations of the Cambridge database paper would also not sufficiently capture the level and quality of research supported. Tracing back actual users and research groups when logging on would provide some additional information. The provision of databases is considered too underpinning to expect research quality to be measured directly.

The level of response from the e-mail survey initiated by the service was considered poor (only 195 responses from 3621 e-mails sent). The low response does not provide a positive indication of the importance of the service in delivering high quality research.

Grade: No Grade

2005 Grade: Red

2.5 Management

There is a high level of staffing, it is unclear as to why such a level is required; however, there are labour intensive elements such as the publicity programme. The information provided on staffing remains unclear, the categories used do not provide an indication as to what the roles actually involve. The summary of staff roles provided in appendix A appears to be that originally presented in January 04. This information should be clarified to include an indication as to what the categories listed against individuals mean in practice; for example is user support primarily delivered through e-mail, telephone or site visits? It would be useful to breakdown individual workloads by particular databases and staff effort in trialing of new databases should also be clarified. The MAP and management team should work together in developing a format to clarify staff roles.

Grade: Red

2005 Grade: Amber

2.6 Performance

The service was clearly working and delivering access to databases; however, the panel considered that insufficient detail was provided on the internal operation of the service and the role of the staff members to be able to assess this aspect. Greater clarity was requested in the previous cross service report; however, this has not been provided in the current annual report.

The performance benchmarks and statistics provided by the service were difficult to understand and require clarification. Performance benchmarks should be put in place that relate to delivering of elements in addition to access of data such as training. The staff effort required to ensure delivering of data from the individual databases was unclear, particularly with regard to the Cambridge database which is accessed using their own proprietary software.

Grade: Red

2005 Grade: Amber

2.7 Training

There has been a change in emphasis within the road shows to include elements on training; however, it is not clear whether training was carried out on all road shows. The service has been directed in previous reviews and the inclusion of these training elements should be seen as a step in the right direction. The panel considered that on all road show visits there should be an element of hands on training. There had been progress but such training could be developed further.

Grade: Amber

2005 Grade: Amber

2.8 Publicity

Major efforts have been made in this area based on the recommendations of the 2003 review panel. The panel acknowledged that the service has been an exemplar in carrying out a coherent and hands on publicity activity. The panel was also impressed that the service had proactively gone beyond chemistry departments to visit biology and chemical engineering departments. This has led to new users being engaged with the service; however, not necessarily at the level expected with certain databases not experiencing the increase in usage expected. There has been a 35% increase in uptake in users from chemical engineering departments.

The impact of the publicity effort appeared to be positive only in part. The major increase in new users was for the structural and DETHERM databases. The renewal panel's original recommendations had been based around raising the awareness of the chemistry community of the whole portfolio of databases including the synthetic organic and other physical databases; the impact of the publicity on these databases has been much more limited.

Grade: Green

2005 Grade: Green

2.9 Overall

The service is competing against commercial database providers, and the advantages of using many of the databases provided by this service are not obvious. There continues to be patchy use of the portfolio of data available, with the structural and DETHERM databases appearing to be the only highly accessed databases. The advantage of the service provision versus commercial provision of the data is not clear. The costs of the national licence for DETHERM is quoted at €173k, if on-line searches had been used this would have been €220k. The overheads associated with service provision make the cost reduction negligible. In addition there are other bodies in a position to negotiate a regional or national licence at a reduced cost.

The added value of the service format is unclear; the primary rationale for the service appears to be negotiation of reduced cost national licences. While this is valuable, it should not be a primary rationale for a chemistry national service. The service does provide other roles such as training and a single portal for entry, though the data providers could provide the former at a certain level and the latter is not a significant advantage for the majority of users who focus on a small number of databases.

The service appears to operate adequately but the panel continued to be unclear as to why a service continued to require significant staff resource. The panel continued to be impressed by the service efforts on their publicity.

The documentation provided by the service while comprehensive was not clear in highlighting the critical information. A significant amount of the data could be summarised pictorially and some of the crucial information was included in an annex, which had to be accessed on line.

Overall Grade: Red

2005 Grade: Amber

3 Requirement for a Database Service for UK Chemistry

3.1 Introduction

In order to separate the question of whether a service is required from the assessment of individual proposals, the cross service panel also carried out the initial assessment of whether the service continues to meet the criteria of a Chemistry National Service in the medium to long term. In order to facilitate this, additional information was provided including contextual data on EPSRC funding and also a vision statement from the current service provider.

A vision statement from the CDS was requested. As the Chemistry Programme has funded existing services for a number of years, the CDS at this initial stage were asked to highlight the case as to why a service is still required to support internationally leading research and how their vision of a service will best meet user needs. The document should outline the continuation of user need, including any areas of chemistry research where the technique is currently under-utilised and the service could raise awareness.

The panel were also provided with an overview of Chemistry Programme, which provided a breakdown of the areas and type of support. Prior to the meeting, the CDS annual report and vision statement was also sent out to a technical assessor.

3.2 Initial panel comments

The panel examined the service against the key chemistry national service criteria; case for a service, uniqueness, cost effectiveness, demand, research quality, and training requirement. As an initial starting point to their discussions the panel was asked to summarise the primary pluses and concerns regarding a chemical database service.

Pluses

- The service is an important resource for a significant number of structural chemists
- “Free” access to data for elements of the chemistry community
- Able to negotiate costs reduction for national licences.
- Able to promote use of specific databases to communities.
- Single point of advice for a number of databases

Concerns

- What is their role in an “information rich” era
- How valuable are the resources currently provided by the service to the whole chemistry community?
- Other than data provision are the other elements of the service of significant value?
- Service model not appropriate with other methods of data provision.
- Key databases are effectively provided through other mechanisms
- What is the added value of a service versus simply obtaining a national licence?
- Can a single service keep up to date with the data requirements and provision across the chemistry community?

The discussion was subsequently structured as a detailed examination against the criteria.

3.3 Case for a service

The panel agreed with the requirement for the UK chemistry community to have access to high quality data. The provision of such data has changed radically over the past 10 years and continues to do so with the developments in GRID and e-science. There is also the increasing use of institutional repositories for data. It is unclear how the national service model, which was developed over 10 years ago, continues to be the most appropriate for the provision of data across its portfolio. The service appears to be focussing on the smaller / specialist databases with major key databases provided through other routes; however, a national service should be central to the provision of a particular technique in underpinning the chemistry community. The panel could not see the case for a service based on small specialist databases, the cost effectiveness of a national licence would be less compelling with specialist databases with limited numbers of users.

3.4 Uniqueness

The service was not considered unique, as academics were able to access data through a number of different avenues. A number of databases that are currently in the service portfolio are also available elsewhere, with some institutions having local access for a number of databases. Other than cost there appears to be little reason why individual researchers and institutions cannot access all their data requirements through direct contact with the vendors. There is an obvious case for cost effectiveness of national licences and this will be discussed below.

The interfaces provided by the vendors are improving and while a key component of the service used to be provision of user-friendly interfaces this appears to be less of the case at present. The service is potentially unique in offering a central portal and the ability to automate the searching of data across a number of databases; however, with communities such as organic and physical chemists using data provided outside the service then this is of less use.

3.5 Cost Effectiveness

A key element of the current service is the negotiation of a national licence for specific databases. Figures are quoted by the service for individual licences; however, it is unlikely that researchers or institutions would continue to purchase this alone. The negotiation of a major regional or national licence could be carried out by another body such as JISC / Eduserv. In addition there are a number of regional clusters such as ScotChem and N8 who would be in a position to examine regional provision.

3.6 Research Quality

Significant efforts have been made in generating exemplars of the research supported by the service. However, as the elements provided by the service are underpinning it is difficult to make any form of accurate assessment of research quality. Based on the information provided by the service (primarily highlights and publications) the quality is not sufficiently high to justify the service. However, the panel considered that especially for the structural databases some high quality research was being supported but in such an underpinning way it is difficult to measure. The assessor comment provided an appropriate analogy with utilities and measuring the quality of research they supported.

The landscape for funding such underpinning elements has also changed in the light of full economic costing, where elements such as database provision should be included in all relevant research grants as part of the infrastructural support.

3.7 Demand

At the time of the last renewal a major effort was established to publicise the service more effectively. The panel were concerned by the relatively low use of the service outside of the structural databases. The publicity effort was focussed on broadening the user base and also highlighting the provision of DETHERM within the portfolio. The service did initiate a questionnaire study across nearly 4000 of its users. The response was considered poor and did not provide a positive reflection of the demand for the service.

Subsequent to this renewal the service has been assessed annually along with the other chemistry national services. The health check continued to raise concerns that the organic databases and the physical ones, other than DETHERM, were still under utilised. In the latest figures provided by the service, the publicity campaign appears to be most effective in increasing the usage of the structural databases and DETHERM. The figures for the other databases continue to appear relatively low especially as there was greater capacity for increased usage of these other databases. Having carried out an effective publicity campaign visiting all the key institutions it would appear that the demand for these other databases is not sufficient to warrant their provision through a national service.

The case for a national service rests primarily on the usage of the structural and DETHERM databases.

3.8 Training

Service training is primarily delivered through the website, though some hands on training is also provided. The increasing provision of usable interfaces provided by the vendors and the opportunity to contact the data providers directly, makes the case for a training programme less clear. An example of this is the Cambridge Crystallographic Data Centre, which has user support along with a page of FAQ's and downloads. As this is the key database within the portfolio, the training programme which is an important rationale behind a service, is not clear in this case.

3.9 Conclusions

The panel concluded that the chemical database service no longer met the criteria of a chemistry national service. It did not appear that the service model was supporting the breadth of the user base and it was difficult to demonstrate that high quality research was being supported across the database portfolio. The centralised format and existing expertise did not appear the most suitable for effective data provision across the chemistry community, with major sources of data such as SciFinder being sourced independently. The additional elements provided by a service specifically training and user support were also offered by the providers independently.

The service had made significant efforts in broadening the user base; however from the data provided a similar spectrum of usage exists at present as in 2003. Most significantly the level of usage of synthetic organic community had not shown a substantial level of increase. The only two key elements of data provision, which continue to have significant usage, are the structural and DETHERM databases.

The service vision aimed at developing a portfolio of specialist databases; a national service should be providing a core technique or in this case key databases. The existing smaller databases do not have significant usage and it is not clear how increasing the portfolio of specialist databases will change this. The developments in data-provision via GRID / e-science mean that a centralised service no longer appears appropriate.

The main area where the service is adding value is in providing a national provision for the two key databases. This is not a major argument for retaining the service, as there are other bodies which should be able to carry out the negotiation of national licences. In addition, there is now co-operation of institutions on a regional level and more specialist databases could be provided regionally. There is now also the opportunity of recouping such funding via infrastructure support on research grants.

The panel was unanimous in their view that a national provision for the structural databases (the Cambridge and inorganic structural databases) and DETHERM remained. EPSRC should explore the requirements of providing such national provision and look to put elements in place to ensure that community access is not disrupted following the end of the current service grant. The main area of uncertainty is whether there is a requirement to physically host and manage the data by a third party or whether users can simply gain access direct with the provider.

Annex 1 Panel membership and Terms of reference

Panel Terms of Reference

- To read and reflect on the papers in advance of the panel meeting
- To fully participate in the discussion at the panel meeting
- To treat all papers and information obtained during the visit as 'in confidence' until final conclusions have been made and reported to the applicants
- To divulge any conflicts of interest

Panel Membership

Prof David Jackson (Glasgow) – Chair
Dr Malcolm Halcrow (Leeds)
Dr Philip Lightfoot (St Andrews)
Dr Jim Thomas (Sheffield)
Dr Richard Wheatley (Nottingham)
Dr Michael Whittlesey (Bath)
Prof Chris Willis (Bristol)

Annex 2 List of Services and Directors

Service	Institution	Director
Mass Spectrometry	Swansea	Prof Gareth Brenton
X-ray Crystallography –	Southampton	Prof Mike Hursthouse
Synchrotron service element	Newcastle	Prof Bill Clegg
Electron Paramagnetic Resonance	Manchester	Dr David Collinson Dr Eric Mcinnes
Solid State NMR	Durham	Dr Paul Hodgkinson
Chemical Database Service	CCLRC	Prof Martyn Guest
Computational Chemistry	Imperial	Prof John Dyke