CCLRC DARESBURY LABORATORY
EPSRC NATIONAL CHEMICAL DATABASE SERVICE

APPENDICES TO ANNUAL REPORT DOCUMENTS 2004/5

1. Competencies of Current Support Staff of the Chemical Database Service
2. Mapping of time spend by individual Staff Members
3. Protocol for Termination of Database/Software Support
4. Protocol for Appraising and Acquiring New Databases and Systems
5. Management Advisory Panel
6. Complete List of Roadshow Visits
7. CDS Roadshow new user survey - January 2005
8. CDS Roadshow Lecture Feedback
9. Trial of Accelrys Databases
10. Components Available during Report Period
11. CDS Service Levels
12. CDS Usage Statistics
13. CDS Usage and RAE Grade for Chemistry Departments

Appendix 1

Competencies of Current Support Staff of the Chemical Database Service

A mapping of time involved for the individual staff members with respect to core CDS functions is given in *Appendix 2*. Some clarification of the skills involved for these functions is given below.

1. **User support**: This includes answering queries, registering users and the preparation of relevant sections of the website. It requires knowledge of the various service components, relevant chemistry disciplines, client computer systems and web authoring skills.
2. **Training**: This includes delivering courses and preparing online material. It requires detailed knowledge of the appropriate service component and chemistry discipline and web authoring skills.
3. **Publicity**: This includes site visits, mailshots, newsletters, the user meeting and web-based material. It requires knowledge of the various service components and relevant chemistry disciplines as well as web authoring, communication and presentation skills.
4. **Database administration and maintenance**: This includes database updates, website maintenance, the installation of fixes and of new releases. It requires detailed knowledge of the database systems, scripting, programming and Oracle DBA skills.
5. **System administration and development**: This includes keeping the systems secure, running and healthy as well as improving the system to user interface. It requires system administration, scripting and programming skills as well as technical knowledge of the various CDS packages and utilities and the client systems which interact with them.
6. **Database development**: This includes the installation of new databases, the building and deployment of new interfaces, the installation and configuration of new hardware as well as the improvement of the CrystalWeb interface and the ongoing migration away from legacy code. It requires detailed technical knowledge of the relevant packages and legacy code, scripting, programming and system administration skills. It also requires knowledge of the current state of chemical information systems and user requirements.
7. **Strategic development**: This includes maintaining a watching brief on developments, attendance at conferences, collecting user requirements and planning. It requires knowledge of the current state and of user requirements, contacts with vendors and others in the chemical information field, communication and negotiation skills.
8. **Service management**: This includes regular reports, MAP meetings, financial management and other administrative tasks.

Appendix 2

Mapping of time spend by individual Staff Members

Function	**"Ideal" case estimate (FTE)**	RFM	DP	DAF	DO-E	**Total (FTE)**
User support	**0.75**	5%	25%	10%	25%	**0.65**
Training	**0.5**	5%	20%	5%	10%	**0.4**
Publicity	**1**	5%	30%	10%	45%	**0.9**
Database administration and maintenance	**0.8**	20%	10%	20%	10%	**0.6**
System administration and development	**0.4**	5%	-	20%	5%	**0.3**
Database development	**1**	40%	5%	30%	5%	**0.8**
Strategic development	**0.25**	10%	5%	5%	-	**0.2**
Service management	**0.15**	10%	5%	-	-	**0.15**

The "Ideal" case corresponds to the resources we would hope to be able commit to the various individual tasks in the absence of the inevitable constraints related to available manpower

Appendix 3

Protocol for Termination of Database/Software Support

*1***.** The full database/utilities portfolio will be formally reviewed periodically by the Service:

a) Every 12 months

b) When there is a major system change (e.g. transfer to new server)

c) Individual items may be reviewed on an *ad hoc* basis where appropriate.

2. Candidates for removal will fall into the categories

a) Limited Usage

- Less than on average 5 accesses per month or low usage out of balance with the ongoing support costs

b) System Superseded

- Better system available on the CDS (e.g. various file conversion utilities now superseded by BEDLAM; FNMR data accessible via SpecInfo)

c) Readily available Elsewhere

- Other Services/Sites provide ready access (e.g. PDB protein datafiles on EBI, and other web sites)
- Software can be downloaded and run more effectively on users' localworkstation (e.g. molecular display packages such as RasMol, the crystallographic package, PLATON)

d) Not Maintainable

- System requires hardware and/or software which is now redundant – would require excessive effort for continued support (e.g. in the future this will increasingly apply to the various legacy codes – indeed certain functionality already dropped for CSSR)

3. Where a specific item meets one or more of the removal criteria it will be examined in more detail. Any contentious cases will be sent to the MAP for discussion.

4. Where it has been determined that an item will be removed

a) Specific active users of the system will be contacted where relevant

b) References to item will be removed from relevant documentation Web Site, Promotional Material, Reference Sheets, Online Help, etc.

c) Notice will be given to the general user community by E-mail and/or the CDS Newsletter where relevant

5. Access to database/utility by users will be barred after completion of the above procedures

6. Total Removal from the system will occur

a) After 6 months

b) If there would be a major support overhead in restoring system to usability

Appendix 4

Protocol for Appraising and Acquiring New Databases and Systems

1. Confirm that the prospective new database contains high quality data that would be of value to a reasonably large number of CDS users at reasonable cost. Criteria for deciding this will include:

> *a)* Prospective systems regarded as of high potential value if they occur frequently amongst suggestions for new data in responses to CDS surveys of user requirements.
> *b)* The quality of new data is likely to be guaranteed if it is supplied by known and reliable source. Such sources currently include: MDL, Accelrys, CCDC, NIST/FIZ, Tripos, Daylight, etc.
> *c)* Where a prospective source is new or not well known the Service will solicit advice from the MAP, other experts, discussion lists, etc. In some cases a specific user survey will be conducted under Item *2*. Particular note will be made of the tests for quality and usefulness under Items *5* & *6.*

2. Inform the CDS MAP of the potentially valuable database. Add "new databases" as a standard item on agenda for MAP meeting.

3. Assuming that the necessary systems are available, attempt to obtain a free trial of the database for a period of several months.

4. Inform the user community of the trial, inviting them to try out the database.

5. Record usage of the database during the trial (as a measure of the community's interest).

6. Toward the end of the trial, evaluate the user community's perception of the database via a survey. This will encompass both the number of users who want the database to be available as well as its value to their research.

7. Submit a proposal for obtaining the database to the CDS MAP, including details of funding.

8. If all the above conditions have been met, obtain the database and release it to the community.

Appendix 5

Management Advisory Panel

Dr. K.A. Johnson, University of Liverpool (retired December 2004)

Dr. A.J. Blake, University of Nottingham (retired December 2004)

Dr. Wendy A. Warr, Wendy Warr & Associates (retired December 2004)

Dr. A. Whiting, University of Durham (retired December 2004)

Prof. R.E. Hubbard, University of York

Dr. N. Greeves, University of Liverpool

Dr. Jeremy Frey, University of Southampton

Dr. William (Bill) G. Town, Kilmorie Consultants

Dr. Mark J. Biggs, University of Edinburgh

Prof. Chick C. Wilson, University of Glasgow

1. The Management Advisory Panel (MAP) exists to assist the Service Director in the effective operation of the National Service by:

a) Ensuring Service is fully utilised in supporting the highest quality science.
b) Advising on the special and changing research needs of the communities using the Service and how the Service might be developed to meet these needs.
c) Advising on how the Service is perceived both scientifically and organisationally by its user communities so that timely action may be taken to build on strengths and address weaknesses.
d) Assisting in the promotion of the Service to ensure that as many as possible of the researchers who might benefit from it are aware of its existence and technical capabilities.

2. The full MAP will normally meet twice a year, and members may be asked to participate in additional meetings involving users or EPSRC as necessary.

3. Members of the MAP should declare any personal interests and not participate in discussions where there would be a conflict of interest.

4. Membership of the Management Advisory Panel should ensure that the MAP has representatives from each main user community and should be reviewed on an annual basis to reflect changes in the user base.

5. The composition of the MAP will be based on the following criteria:

a) MAP membership should be for a fixed three year term.
b) New MAP members to be chosen in consultation with EPSRC.
c) MAP to have at least one member from industry.
d) MAP to have at least one member who coordinates chemical information provision at a major UK university.
e) MAP to include no more than one member from any single institution.

It should be noted that in the *Terms of Reference* listed above item *5(d)* has been changed. During discussions on *Grant Renewal Issues* it was agreed that the MAP should include "a member from the UK academic library community". This was changed as a result of recommendation form the MAP itself meeting 20/12/04. It was felt that coordinating access to chemical information was increasingly being devolved to Departments. The current MAP already includes members with involvement in this process in their Departments.

Appendix 6
Complete List of Roadshow Visits

Date	Site	1	2	3	4	5	6	7	8
29/10/03	Liverpool Chemistry	35	35		6	5			
30/10/03	Manchester Chemistry	25	50		11	3			
26/11/03	Loughborough Chemistry	50	30	15	14	1			
20/4/04	Sussex Chemistry	12	10	9	6	2			
8/6/04	Imperial College Chemical Engineering	22	25	15	8	2			
30/9/04	Oxford Chemistry	17	24	12	14	2	69	20%	108
4/10/04	Warwick Chemistry	30	22	50	22	8	56	39%	78
6/10/04	UCL Chemistry	10	25	18	6	1	55	11%	119
13/10/04	Swansea Chemistry	12	15	12	3	4	3	100%	6
14/10/04	Cardiff Chemistry	60	50	45	28	1	30	93%	41
15/10/04	Nottingham Chemistry	12	28	30	2	1	168	1%	199
21/10/04	Leeds Chemistry/Chemical Engineering	6	15	20	2	6	95	2%	126
26/10/04	Sheffield Chemistry	50	50	30	26	4	55	47%	93
28/10/04	Newcastle Chemical Engineering	2	20	0	1	2	4	25%	34
29/10/04	Durham Chemistry	20	50	20	22	1	84	26%	96
8/11/04	St. Andrews Chemistry	20	20	20	8	2	57	14%	67
9/11/04	Edinburgh Chemistry	30	30	30	26	1	43	60%	103
11/11/04	Glasgow Chemistry	40	33	13	29	0	15	193%	31
12/11/04	Strathclyde Chemistry	6	22	6	1	1	37	3%	51
19/11/04	Bangor Chemistry	25	50	27	21	1	8	263%	11
25/11/04	Liverpool Chemistry	20	27	15	13	1	68	19%	77
30/11/04	Manchester Chemistry	15	47	12	14	15	60	23%	85
25/1/05	Southampton Biological Sciences	5	20	4	2	1	8	25%	213
9/2/05	Bath Chemical Engineering	12	20	12	10	1	1	1000%	77
10/2/05	Bristol Chemistry	15	20	24	12	3	83	14%	101
17/2/05	Loughborough Chemical Engineering	25	20	14	13		2	650%	61
1/3/05	Birmingham Chemistry	5	25	10	6		44	14%	64
8/3/05	Cambridge Chemistry	7	20	7	2		95	2%	206
15/3/05	York Chemistry	5	20	9	5		70	7%	75
22/3/05	Belfast Chemistry	20	33	15	11		24	46%	44
	Totals	**613**	**856**	**494**	**344**	**69**	**1234**	**28%**	**2166**

Column headings:

1: Number of people talked to at the stand
2: Number of CDS leaflets distributed
3: Number of people at the lecture
4: Number of new users signed up at the event
5: Number of new users signed up in the next month
6: Number of users at the Department before the visit
7: Percentage increase in users at the Department due to the visit
8: Total number of users at the University

Appendix 7
CDS Roadshow new user survey - January 2005

The survey contained the following questions; the number of responses for each answer is listed on the left.

1. Which of the following best describe your reasons for only having made limited use of CDS so far? Please check all that apply. (19 responses)

10	A lack of time
4	Problems with accessing the CDS
2	Problems with getting a suitable computer or software to access the CDS
1	Problems using the CDS packages themselves
1	The data available are not relevant to my research
1	The data are not as useful as they first appeared to be
8	I've not needed to access the CDS yet
1	Other data sources are better for my needs
0	Other, please specify

2. What could CDS do or have done in order to improve your usage? Please check all that apply. (13 responses)

1	Had other data available such as
2	Had more online help or tutorials available
1	Provided hands-on training courses
0	Provided different software packages to access the data
0	Other, please specify

3. Did you attend the CDS presentation at your site ? (13 responses)

8	Yes
5	No

4. If you have used any of the CDS website resources, which, if any, did you find useful? (13 responses)

1	Flash based movie demonstrations
3	Overview lecture PowerPoint slides
1	CDS User Guide
4	Online manuals
0	Other, please specify

5. What other electronic chemical information sources do you use regularly? Please check all that apply. (13 responses)

7	SciFinder
5	CrossFire Beilstein/Gmelin
0	NIST (chemistry webbook or other web resource)
10	Web of Science
1	CD-ROMs
0	Other(s), please specify

Appendix 8
CDS Roadshow Lecture Feedback

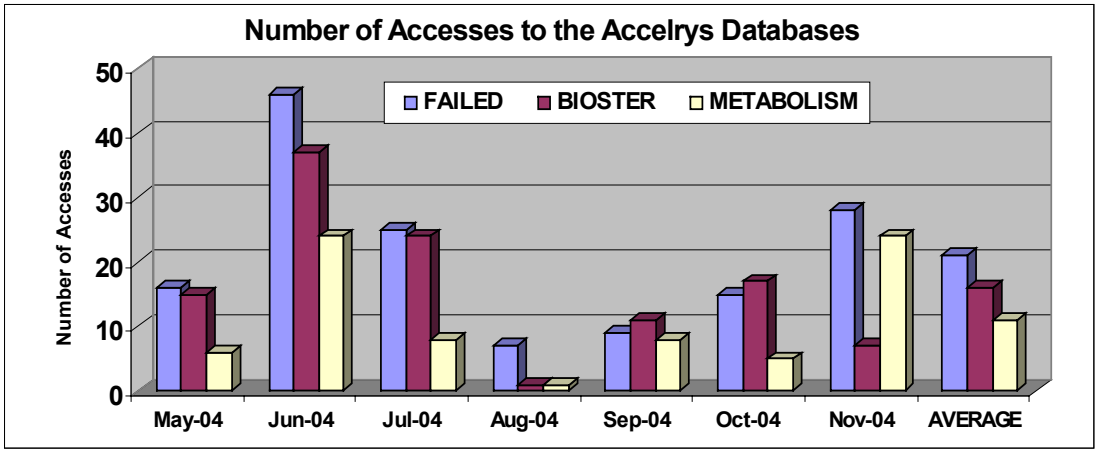
Question	Responses	1	2	3	4	5	6	**Total**	
1	Host	8	6	7	19	7	15	**62**	83.8%
	Other		4	3		5		**12**	16.2%
2	Email	3	3	5	16	7	9	**43**	58.1%
	Poster	5	3	5	1	0	10	**24**	32.4%
	Colleague	3	5	3	3	5	2	**21**	28.4%
	Website	0	0	0	0	0	0	**0**	0.0%
3	Yes	3	2	2	5	2	6	**20**	27.0%
	No	5	8	8	12	10	9	**52**	70.3%
4	Not	0	0	0	1	0	0	**1**	1.4%
	Slightly	1	0	0	7	0	0	**8**	10.8%
	Reasonably	6	7	4	5	8	5	**35**	47.3%
	Very	0	2	5	5	4	8	**24**	32.4%
	Extremely	0	0	1	0	0	2	**3**	4.1%
5	No	1	0	0	1	0	0	**2**	2.7%
	Possibly	1	1	3	5	1	0	**11**	14.9%
	Probably	3	3	2	5	9	3	**25**	33.8%
	Yes	3	6	5	7	2	12	**35**	47.3%
6	CrossFire	6	0	2	18	2	13	**41**	55.4%
	SFS	4	2	6	17	4	0	**33**	44.6%
	NIST	0	0	2	3	1	0	**6**	8.1%
	WoS	7	7	8	13	10	12	**57**	77.0%
	CDROM	2	3	1	1	3	3	**13**	17.6%
	Other	1	1	0	0	1	1	**4**	5.4%
	Total forms	8	10	10	19	12	15	**74**	

Questions (rows):
1: Which Department are you from?
2: How did you hear about today's event?
3: Were you a registered CDS user before today?
4: Was today's talk helpful or useful to you?
5: Do you think that today's visit by the CDS will increase your use of the Service?
6: Other electronic sources of chemical information you use regularly?

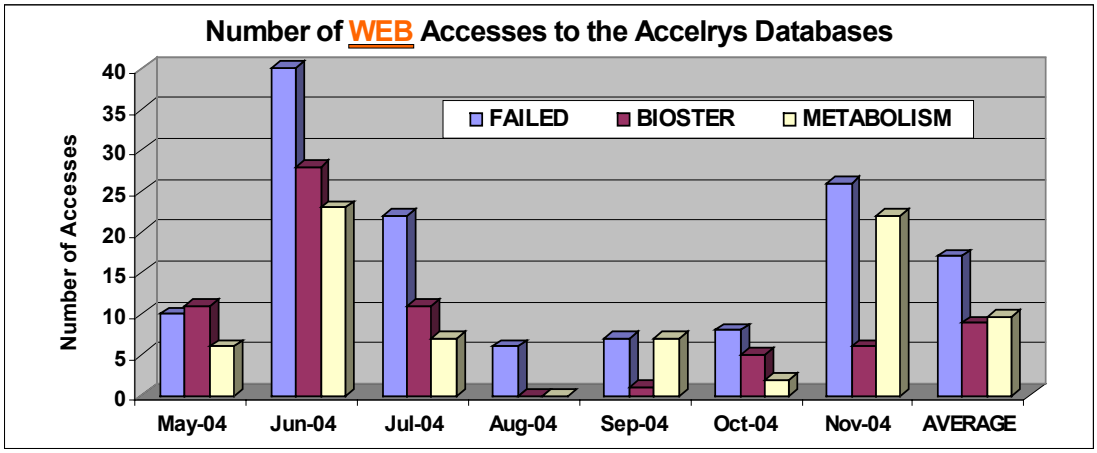
Sites (columns):
1: York Chemistry
2: Loughborough Chemical Engineering
3: Birmingham Chemistry
4: Bristol Chemistry
5: Bath Chemical Engineering
6: Belfast Chemistry

Appendix 9

Trial of Accelrys Databases
FAILED REACTIONS, BIOSTER and METABOLISM

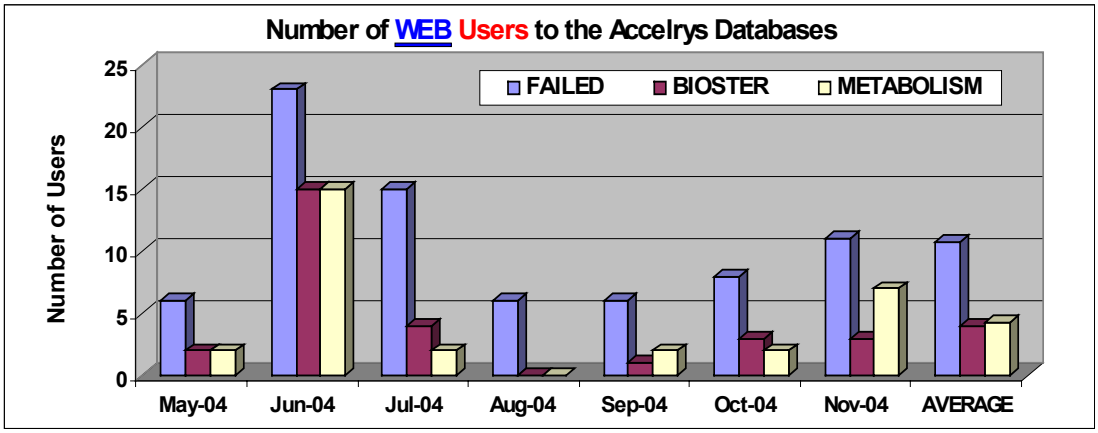


ACCESSES



Web vs Client Accesses	FAILED	BIOSTER	METABOLISM
	81%	55%	87%

WEB USERS



Appendix 9 (cont.)

Database	Have you used this database ?	How useful would this database be in your work ?	Should we acquire this database ?
Failed Reactions	○ Yes ○ No	Select usefulness	Should we buy ?
Metabolism	○ Yes ○ No	Select usefulness	Should we buy ?
Bioster	○ Yes ○ No	Select usefulness	Should we buy ?

Select usefulness: Select usefulness, No use at all, Slightly useful, Sometimes useful, Useful, Very useful, Vital

Should we buy ?: Should we buy ?, No, Possibly, Probably, Yes

Please enter any other comments you have

submit form reset form

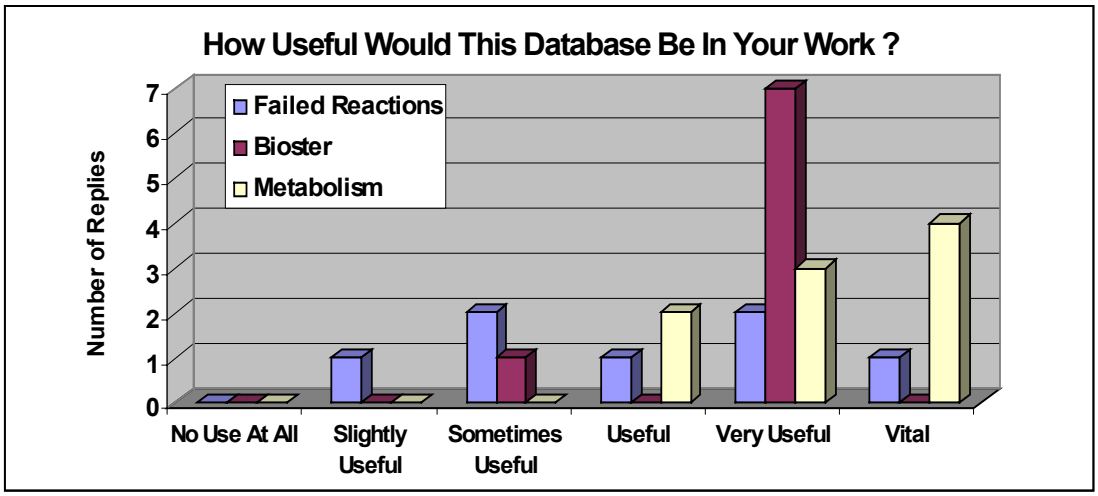
SURVEY

10 Questionnaires filled in

RESULTS OF QUESTIONNAIRE

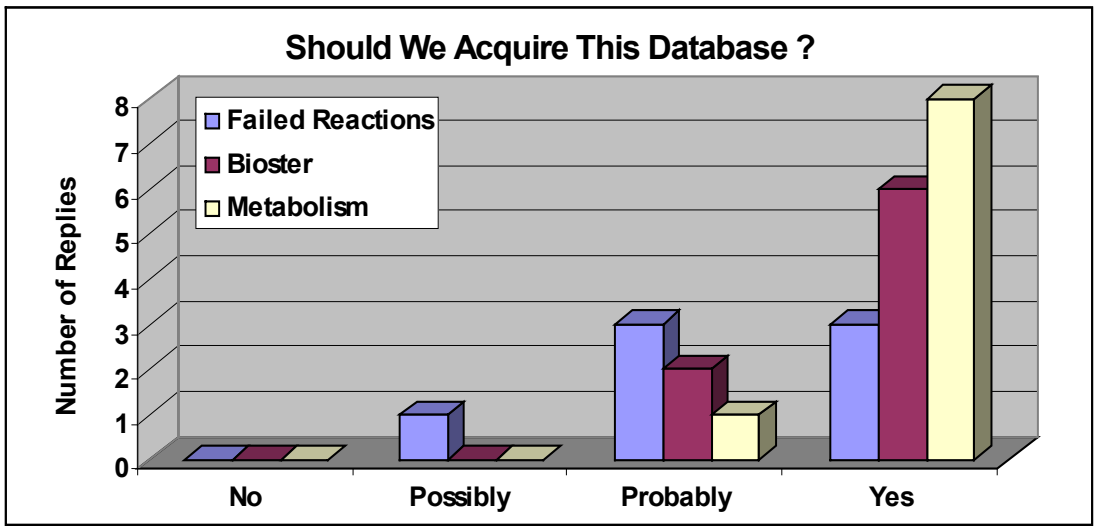
Have you used this database?	YES	NO
Failed Reactions	7	3
Metabolism	9	1
Bioster	7	3

2. How useful would this database be in your work?



Appendix 9 (cont.)

3. Should we acquire this database?



COMMENTS

Metabolism database is an excellent addition to all the other databases CDS currently offer, and I believe it is a vital addition needed. [Dr F. Shah, MRC Clinical Science Centre]

Being not an organic chemist I do not want to judge the Failed Reactions Database; the other two seem to be quite useful. Right now I do not need them for my current research, but they are so close that I would probably use them in the future. [Mr A. Bender, Dept. of Chemistry, University of Cambridge]

The bioster database is a very useful tool when searching for ideas on how to change a core medicinal chemistry scaffold and has been useful to me on a number of occasions. [Dr T. Matthews, Inst. of Cancer Research]

It is very useful to be able to distinguish between reactions that have not been attempted before and reactions that have not been published because they have not worked. It helps assess the risk of a proposed line of research. [Dr M. Christlieb, Dept. of Inorganic Chemistry, University of Oxford]

Appendix 10

Components Available during Report Period

CRYSTALLOGRAPHY

CSD Cambridge Structural Database. Crystal structure data for over 338,000 organic and organo-metallic compounds. 3D geometric search capabilities are available for this data. Accessed via Quest, ConQuest and CSSR.

ICSD Inorganic Crystal Structure Data File. Over 82,000 inorganic structures, searchable via a web interface. Available via the ICSD-WWW web browser interface.

MDF Metals Data File (CRYSTMET). Crystal structure data for over 81,000 metals, alloys and intermetallics.

CDIF Crystal Data Identification File. Crystal class and unit cell data for over 237,600 structures.

All crystallographic databases are accessible via the CrystalWeb interface.

SPECTROSCOPY

SPEC SpecInfo is a multi-technique spectroscopic database system designed to aid the chemist in interpretation and structure elucidation problems. The associated database currently contains 108,000 ^{13}C, ^{15}N, ^{19}F and ^{31}P NMR spectra, 61,000 ^{1}H NMR spectra, 21,000 infra-red spectra and 130,000 mass spectra.

SpecInfo was available by both the established X-Windows interface and SpecSurf, a web browser based interface. The X-Windows interface was withdrawn February 2004.

PHYSICAL CHEMISTRY

DETHERM One of the world's largest collections of thermophysical properties databases of pure components and compound mixtures. It contains over 4.2 Million datasets for over 118,000 systems (24,731 pure substances and 93,505 mixtures) covering more than 200 properties.

The full DETHERM service was made available on 1/4/04.

SYNTHETIC ORGANIC CHEMISTRY

ISIS Chemical reaction information management system allowing search, retrieval and display of molecules, reactions and their associated data. Currently ISIS can access around 1.3M searchable reactions from the following databases:

REFLIB (Reference Library of established literature)
DERWENT-JSM (Journal of Synthetic Methods)
CHEMINFORM (Current awareness database - updated every 6 months)
ORGSYN (Organic Synthesis)
SPG (Synopsys Protecting Groups)
SPS (Solid Phase Synthesis)
BioCatalysis (Biomolecules as catalysts)
ChirBase (Chiral Separations by Chromatography)
NCI (National Cancer Institute Database.)
SCD (Screening Compounds Database - over 3 Million compounds)

ISIS also allows access to ACD (Available Chemicals Directory), which is a database of suppliers of chemicals that contains around 364,000 unique compounds from 670 different suppliers and SCD (Screening Compounds Database) that contains over 3 Million compounds from around 20 suppliers.
Both client/server and web browser interfaces are available for all ISIS components.

Appendix 11

CDS Service Levels

Any service levels which involve people do not apply during a holiday period such as Christmas - New Year. The computers are left running over this period and can be accessed by users. Any serious faults reported will be investigated, with staff being called out if necessary.

1. Database Services

1. New database releases available to users within one working week of reception 95% of the time. Any failures to meet this service level will be recorded as a cumulative total number of days and reported (with full details) to the Management Advisory Panel.
2. Advice on use of chemical information systems, not available in the database service, to be given within two working days with 95% availability in a calendar month.

2. Support

1. Chemist available to answer queries during office hours with 95% availability in any calendar month.
2. General computing queries (high priority) responded to within two working hours during office hours with 97% availability in any calendar month.
3. All other computing queries responded to within two working days with 95% availability in any calendar month.
4. Registration of new users complete within one working week with 95% availability within a calendar month.
5. Bugs and errors in online documentation to be corrected within 2 working days of notification 95% of the time.

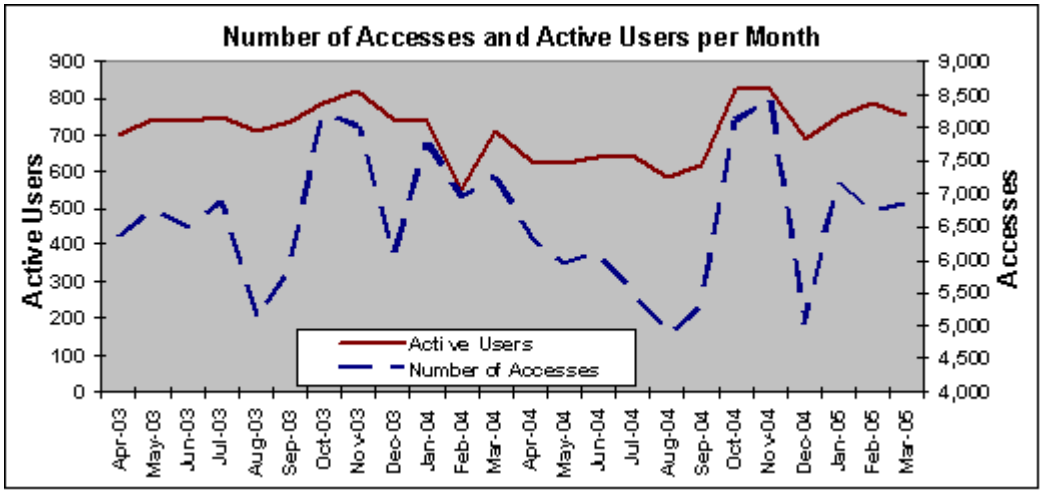
3. System Performance

1. Service availability 99% in any calendar month excluding scheduled down time.
2. Scheduled down time less than 4 hours per quarter.
3. At least two working days warning of scheduled down time via login messages.
4. a) Daily incremental back up of user discs on main machine with 99% successful completion.
   b) Weekly full disc back up on main machine with 99% successful completion.

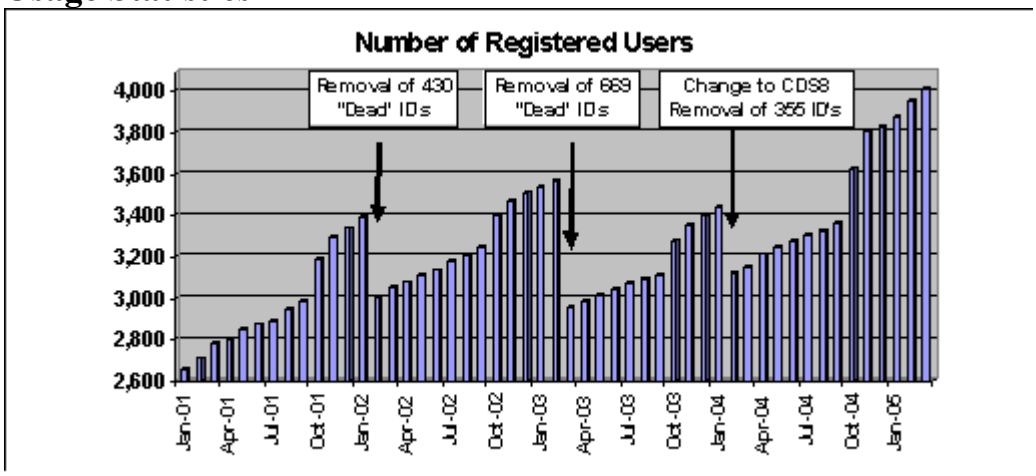
4. Network Access

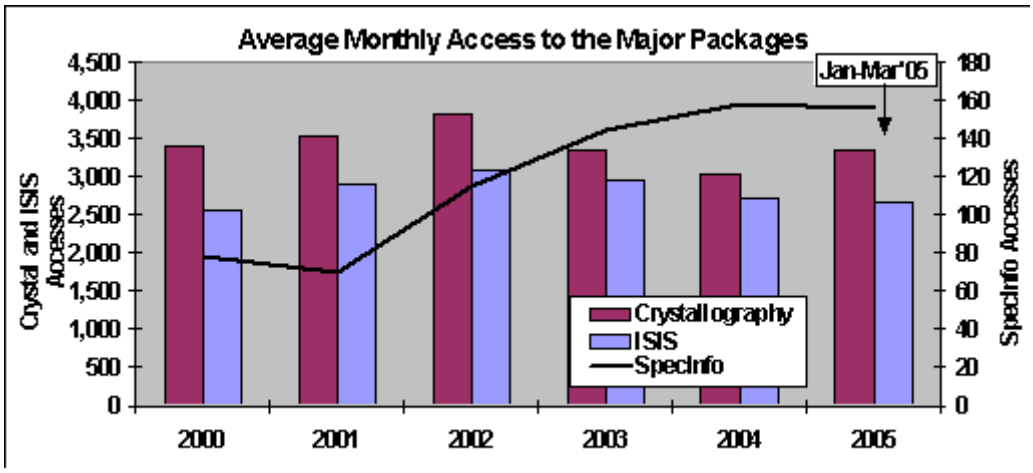
1. 99% Availability of Daresbury campus network (JANET packet switched exchange (JPSE) and campus packet switched exchange (CPSE) and associated on site network) excluding the JNT specified 'at risk' periods (when scheduled maintenance may occur) which are confined to Tuesdays from 8 am to 10 am.
2. We can not guarantee the availability of the academic network (JANET) but will give assistance in tracing network access difficulties within one working day during office hours with 95% availability in any calendar month.

Appendix 12



CDS Usage Statistics





Appendix 13

CDS Usage and RAE Grade for *Chemistry Departments* (April 2004 - March 2005)

Institution CHEMISTRY Department	RAE Grade	Acesses	Registered Users	Active Users	Accesses per Reg User	Accesses per Active User	%Active vs Reg.User
University of Cambridge	5*	3962	97	54	40.8	73.4	55.7%
University of Oxford	5*	2398	86	49	27.9	48.9	57.0%
Imperial College, London	5*	1969	106	26	18.6	75.7	24.5%
University College London	5*	1427	67	44	21.3	32.4	65.7%
University of Durham	5*	1068	105	41	10.2	26.0	39.0%
University of Bristol	5*	452	98	30	4.6	15.1	30.6%
University of Southampton	5	3496	224	70	15.6	49.9	31.3%
University of St Andrews	5	1997	72	42	27.7	47.5	58.3%
University of Birmingham	5	1922	54	35	35.6	54.9	64.8%
University of Leeds	5	1731	113	40	15.3	43.3	35.4%
University of York	5	1647	76	36	21.7	45.8	47.4%
University of Sussex	5	1595	39	22	40.9	72.5	56.4%
University of Manchester	5	1560	83	34	18.8	45.9	41.0%
University of Warwick	5	1326	92	50	14.4	26.5	54.3%
University of Nottingham	5	1115	172	56	6.5	19.9	32.6%
University of Liverpool	5	1108	81	42	13.7	26.4	51.9%
University of Edinburgh	5	891	62	27	14.4	33.0	43.5%
University of Sheffield	5	548	92	33	6.0	16.6	35.9%
University of East Anglia	5	443	32	8	13.8	55.4	25.0%
University of Reading	4	2811	74	33	38.0	85.2	44.6%
Heriot-Watt University	4	2507	43	24	58.3	104.5	55.8%
University of Exeter	4	1928	36	20	53.6	96.4	55.6%
University of Wales, Cardiff	4	1733	60	31	28.9	55.9	51.7%
University of Bath	4	1318	55	28	24.0	47.1	50.9%
University of Hull	4	1010	74	26	13.6	38.8	35.1%
University of Strathclyde	4	986	48	24	20.5	41.1	50.0%
UMIST	4	732	65	18	11.3	40.7	27.7%
University of Glasgow	4	686	41	21	16.7	32.7	51.2%
Loughborough University	4	491	57	23	8.6	21.3	40.4%
The Queen's University of Belfast	4	210	46	17	4.6	12.4	37.0%
University of Newcastle upon Tyne	4	172	25	7	6.9	24.6	28.0%
University of Leicester	4	138	12	6	11.5	23.0	50.0%
Kings (London)	4	101	14	4	7.2	25.3	28.6%
University of Wales, Swansea	4	13	10	4	1.3	3.3	40.0%

Continued on next page

Appendix 13 Continued