

Adding value to open access research data: the eBank UK Project.

Dr Liz Lyon, Director

UKOLN, University of Bath, UK

OAI4, CERN Geneva, October 2005.



UKOLN is supported by: **JISC** 
MUSEUMS LIBRARIES ARCHIVES



www.ukoln.ac.uk

UKOLN

a centre of expertise in digital information management



Overview

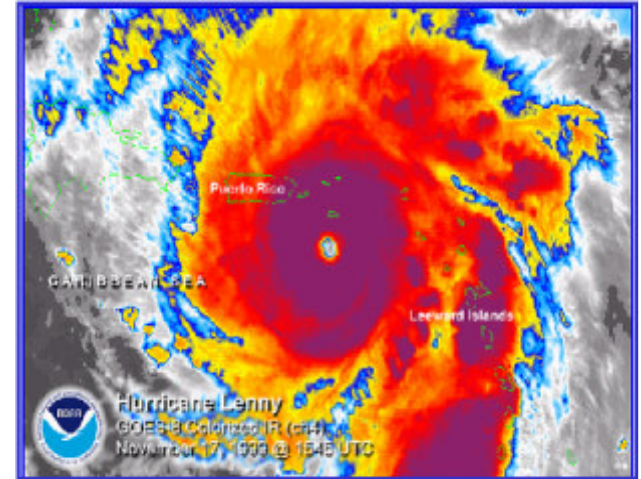
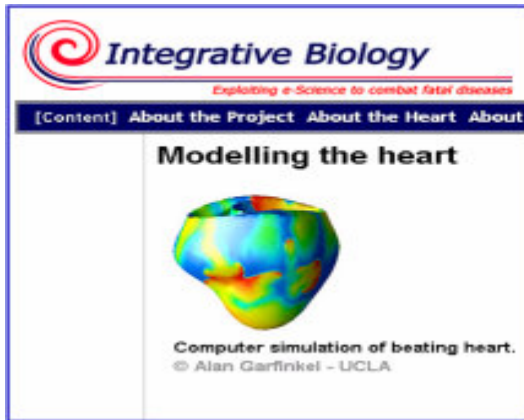
1. e-Research & data-intensive science
2. Repository services & adding value
 - Aggregation and linking: eBank UK
 - Integration and workflows
3. Looking to the longer term: digital curation and preservation



1. e-Research & data-intensive science



eScience - the data deluge



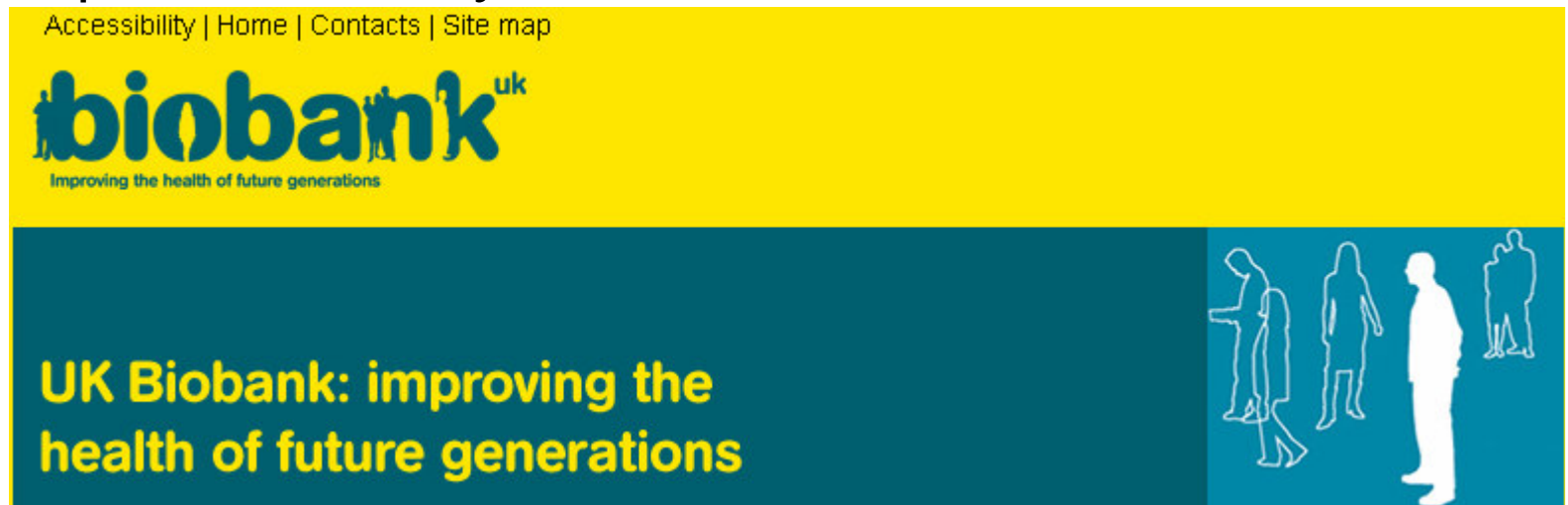
Data
Overload!



How do we
disseminate?

Diversity of data collections

- Very large, relatively homogeneous:
Large-scale Hadron Collider (LHC) outputs from CERN
- Smaller, heterogeneous and richer collections:
World Data Centre for Solar-terrestrial Physics CCLRC
- Small-scale laboratory results:
“jumping robots” project at the University of Bath
- Population survey data: *UK Biobank*



- Highly sensitive, personal data: *patient care records*

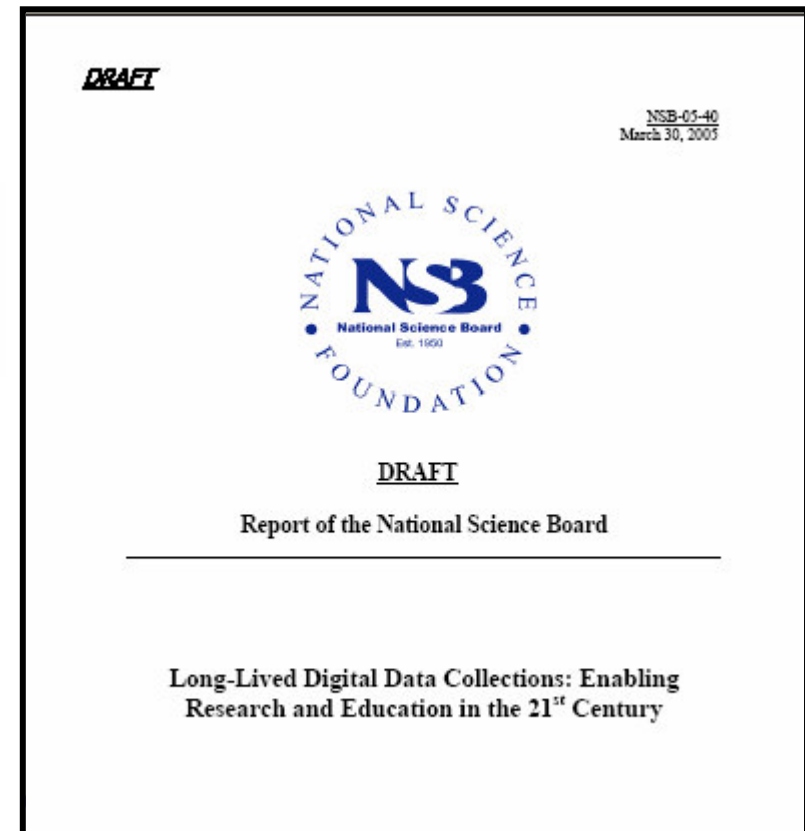
Taxonomy of data collections

- Research collections:
jumping robots
- Community collections:
Flybase at Indiana (with UC Berkeley)
- Reference collections:
Protein Data Bank

RCSB **PDB**
PROTEIN DATA BANK

Evolution.....

FlyBase A Database of the
Drosophila Genome



Source: NSF Long-Lived Digital
Data Collections

Draft report revised May 2005

Experience of data-sharing

- Large scale data sharing in the life sciences
Draft Report June 2005
Sponsored by UK research funding bodies
MRC, BBSRC, NERC, JISC, Wellcome
- Outcomes & recommendations
 - Importance of standards and good quality metadata
 - Require a data management plan
 - Work needed on vocabularies & ontologies
 - Awareness of archiving & long term preservation
- Position of research funders and policy makers?



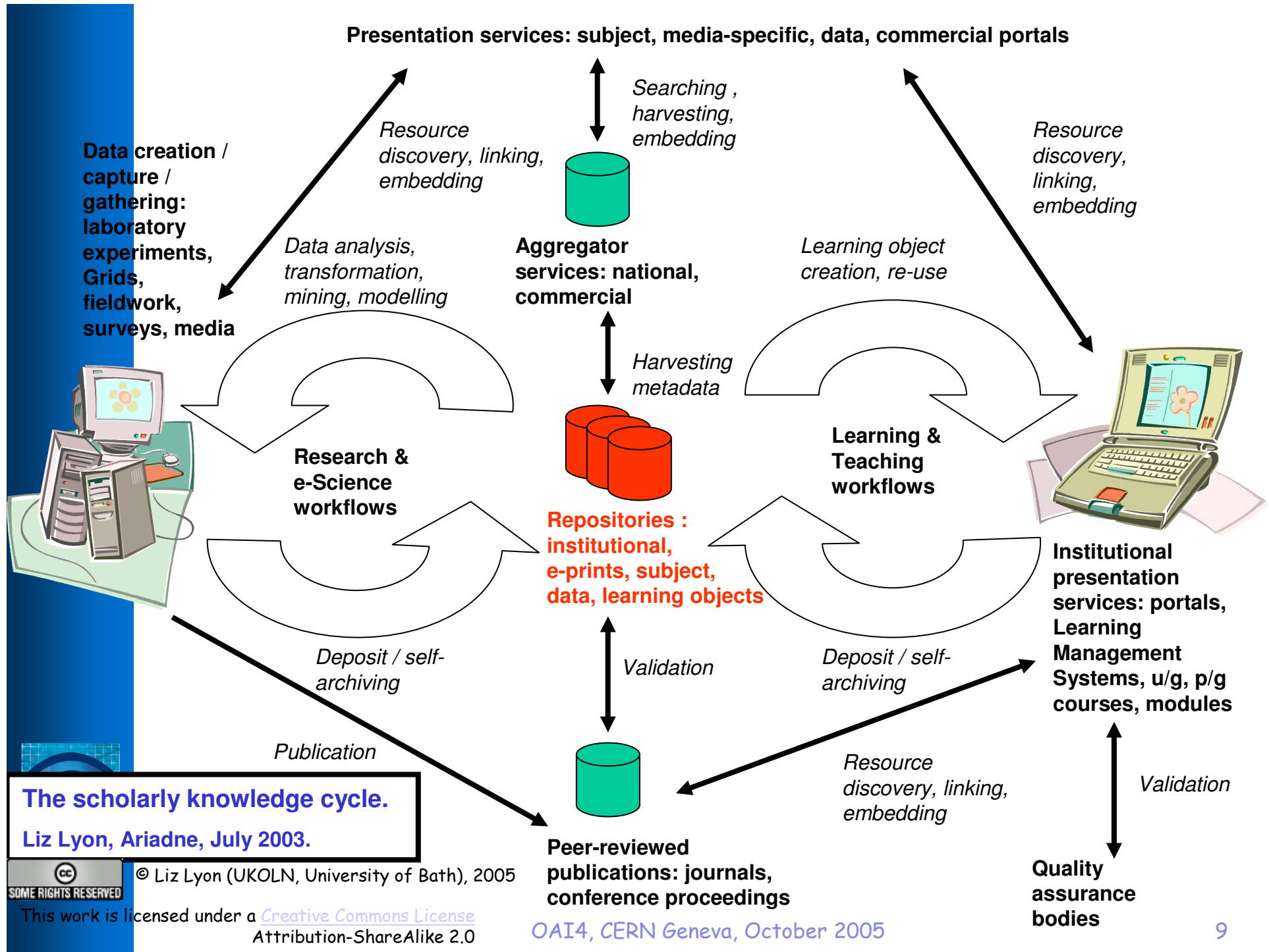
News release

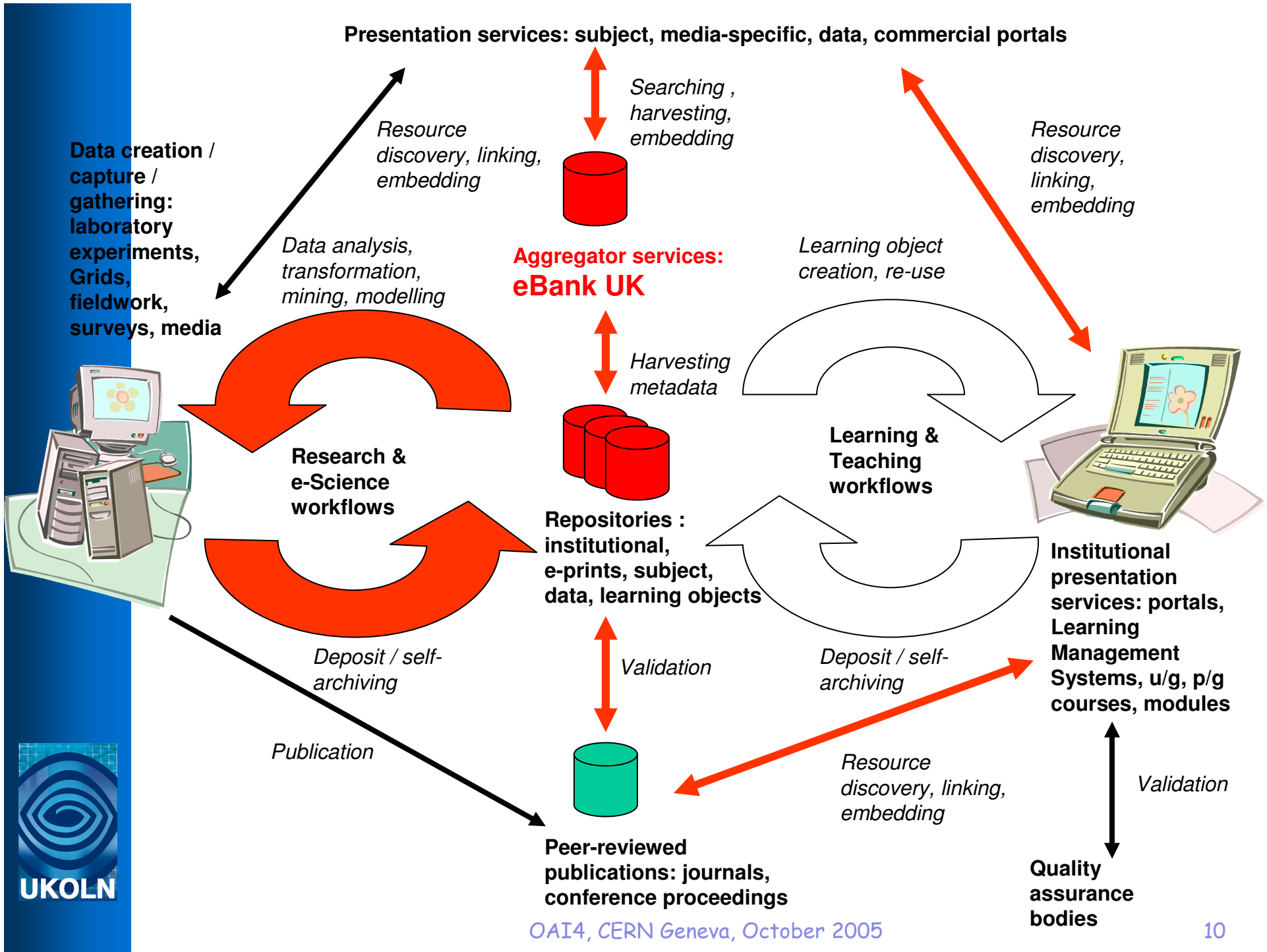
28 June 2005

[RCUK Announces Proposed Position on Access to Research Outputs](#)

Research Data

8. RCUK also notes that one of the benefits of digitisation and publication in digital formats is the ability to provide access to primary research data alongside the traditional article; and it shares the Select Committee's and the Government's view that the data underpinning the published results of publicly-funded research should be made available as widely and rapidly as possible. For a number of years, Research Councils including the AHRB, ESRC and NERC have funded data centres and services which are responsible for preserving, managing and providing access to research data; and these Councils have well-established policies and procedures for preservation and access. CCLRC is currently leading cross-Council consideration of how policy and practice need to be developed with regard to the curation of the data created through the research projects they support. *Further work is needed to develop a common framework of policies and procedures for determining what sets of data are collected, whether in university or in Council-run repositories or elsewhere; and how and on what terms they are made accessible to the research community and others*





2. Repository services & adding value: the eBank UK Project



eBank UK Project

- Two key themes:
 - **Open access to datasets**
 - **Linking research data to publications and to learning**
- JISC-funded from September 2003: now in Phase 2
- UKOLN at the University of Bath (lead), University of Southampton, University of Manchester
- Exemplar: e-Science testbed ‘Combechem’
 - Grid-enabled combinatorial chemistry / crystallography
 - National Crystallography Service
- Resource Discovery Network / PSIgate physical sciences portal
- <http://www.ukoln.ac.uk/projects/ebank-uk/>



The “hybrid” project team



- **UKOLN**
- Michael Day
- Monica Duke
- Rachel Heery
- Traugott Koch
- Liz Lyon
- +
- Andy Powell



University
of Southampton

- **Southampton**
- Les Carr
- Simon Coles
- Jeremy Frey
- Chris Gutteridge
- Mike Hursthouse
- Andrew Milstead

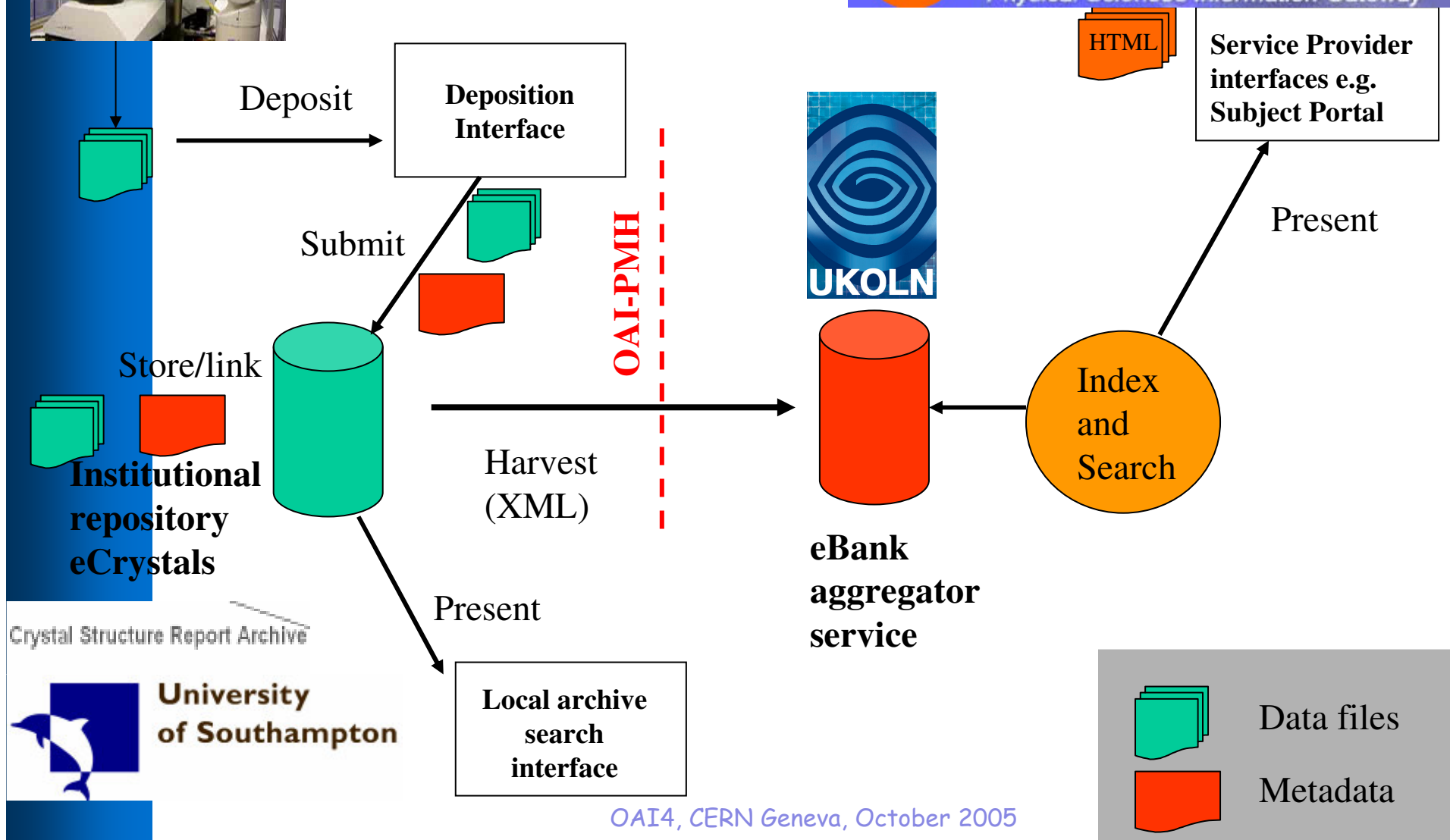
- **Manchester**
- John Blunden-Ellis



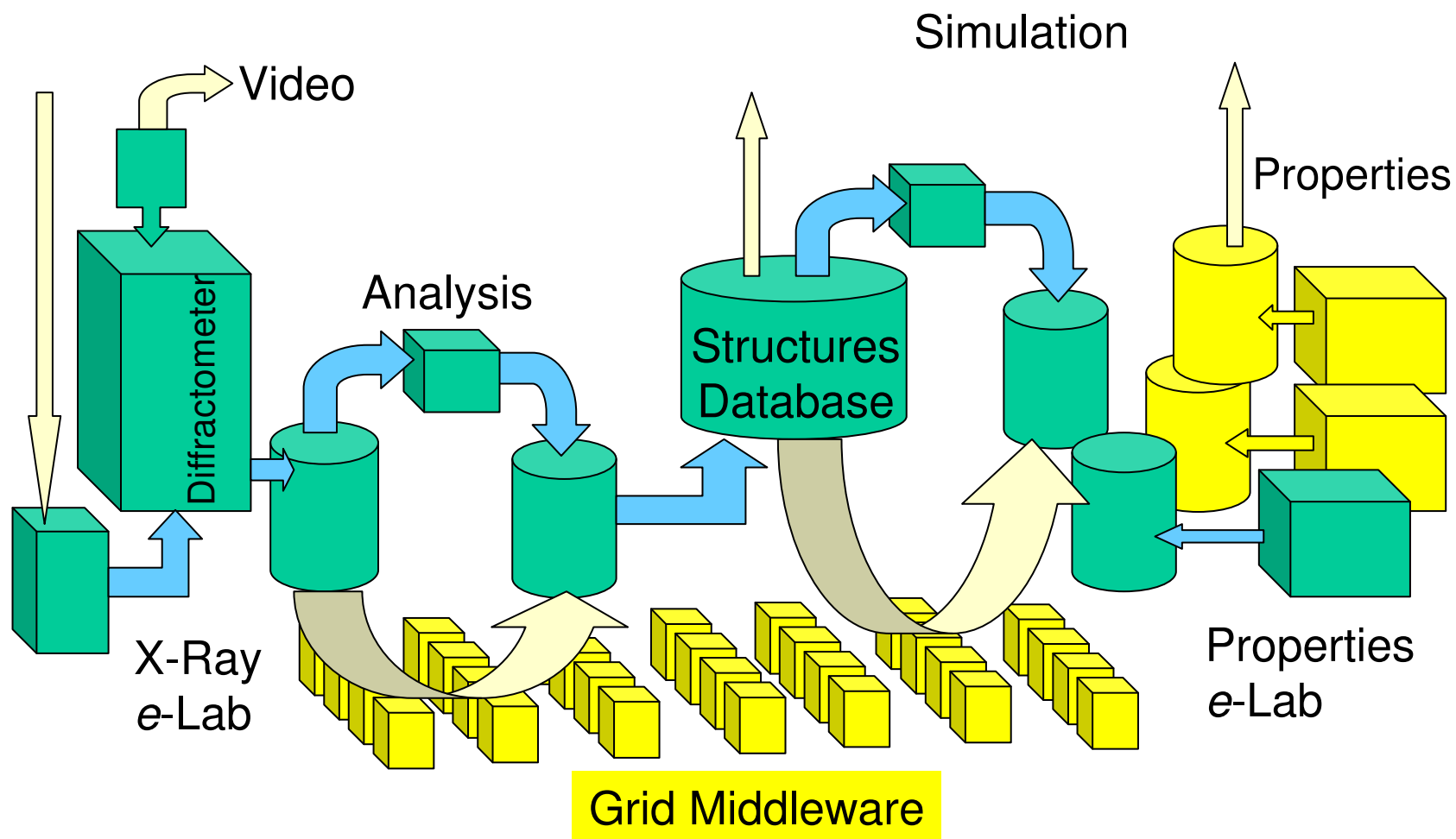
Create



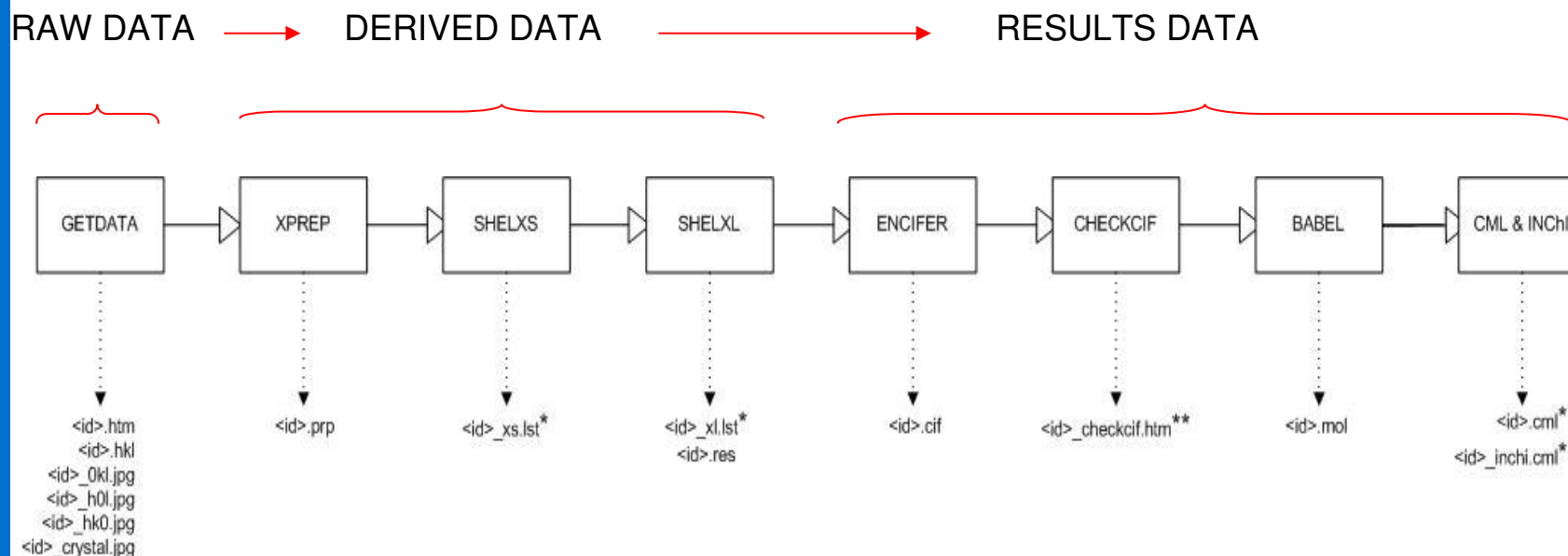
Data Flow in eBank UK



CombeChem: An EPSRC pilot project



Crystallography workflow



- **Initialisation:** mount new sample set up data collection
- **Collection:** collect data
- **Processing:** process and correct images
- **Solution:** solve structures
- **Refinement:** refine structure
- **CIF:** produce CIF (Crystallographic Information File)
- **Validation:** chemical & crystallographic checks
- **Report:** generate Crystal Structure Report



University
of Southampton

Crystal Structure Report Archive

Home

About

Browse

Search

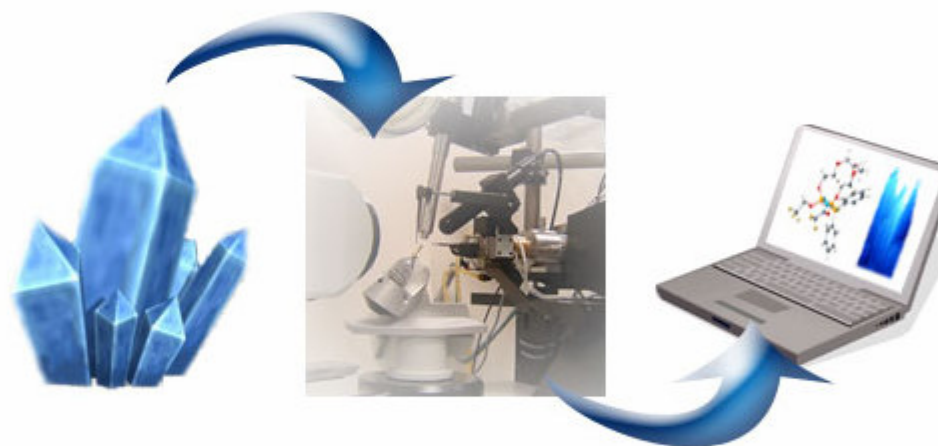
Register

User Area

Help

Welcome to Southampton EBank

Southampton eCrystals is the archive for Crystal Structures generated by the Southampton Chemical Crystallography Group and the EPSRC UK National Crystallography Service.



The information contained within each entry of this archive is all the fundamental and derived data resulting from a single crystal X-ray structure determination, but excluding the raw images. The results have not been externally refereed, but the information supplied should enable any reader to check the reliability and validity directly, since all the files provided are freely available for download. Should any error be detected, we would appreciate receiving suitable comments, and we will make any necessary amendments, and include a note to that effect. Any reader wishing to have access to the raw images is welcome to contact us, and we will make arrangements for these to be made available.



A data repository entry

http://ecrystals.chem.soton.ac.uk/149/

Benzene 1,2dicarboxylic acid

Simon J Coles, Michael B Hursthouse, Claire L Taylor and Peter N Horton.

University of Southampton

C₈H₆O₄

ICHI Code: INChI=1.12Beta/C8H6O4/c9-7(10)5-3-1-2-4-6(5)8(11)12/h1-4H,(H,9,10)(H,11,12) ([google for ichi](#))

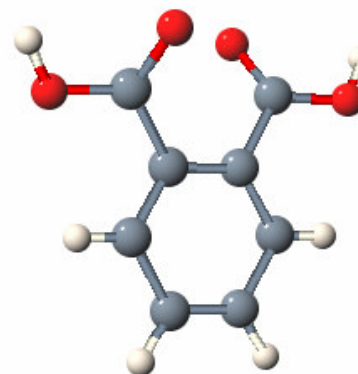
Compound Class: Organic

Keywords: Phthalic acid

Creation Date: 15 February 2005

Deposited By: [Dr Simon J Coles](#)

Deposited On: 21 February 2005



Data collection parameters

Chemical formula	C8 H6 O4
Crystallisation Solvent	
Crystal morphology	Prism
Crystal system	monoclinic
Space group symbol	C2/c
Cell length a	5.0016(10)
Cell length b	14.214(3)
Cell length c	9.5196(19)
Cell angle alpha	90.00
Cell angle beta	94.33(3)
Cell angle gamma	90.00
Data collection temperature	120(2)

Available Files

Final Result

05mbh1006.cml	3k
05mbh1006/05mbh1006.cif	9k
05mbh1006/05mbh1006_checkcif.htm	7k
05mbh1006_inchi.cml	1k

Refinement

05mbh1006/05mbh1006.res	3k
05mbh1006/05mbh1006_xl.lst	21k



UKOLN

Access to the underlying data: complex objects

Available Files

File Name	Size
02sot064.CIF	19k
02sot064.cml	8k
02sot064_checkcif.html	14k
Refinement	
02sot064.RES	9k
Solution	
02sot064.PRP	5k
Processing	
02SOT064.HTM	6k
02sot064.HKL	338k
Other Files	
02sot064.DOC	113k
02sot064.LST	49k

Data collection parameters

Chemical formula	C30 H26 Fe N2 O3
Crystallisation Solvent	
Crystal morphology	
Crystal system	Orthorhombic
Space group symbol	Pbca
Cell length a	6.0816(4)
Cell length b	24.8503(16)
Cell length c	31.120(3)
Cell angle alpha	90.00
Cell angle beta	90.00
Cell angle gamma	90.00
Data collection temperature	120(2)

Refinement results

Solution figure of merit	
R Factor (Obs)	0.0573
R Factor (All)	0.1185
Weighted R Factor (Obs)	0.1046
Weighted R Factor (All)	0.1243

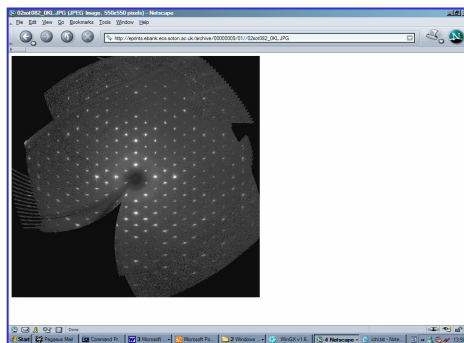
checkCIF/PLATON report (full structural check)

No syntax errors found.
Please wait while processing

Datablock: 1

Bond precision: C-C = 0.0040 Å Wavelength=0.71073
 Cell: a=6.204(3) b=12.261(11) c=10.418(11)
 alpha=90.000(14) beta=90.000(14) gamma=90.000(14)

Volume: 246.24(6)
 Space group: P -1
 Hall group: -P 1
 Z: 2
 Density formula: C12 H8 Cl F N2 O
 Sum formula: C12 H8 Cl F N2 O
 MW: 282.49
 Mo, K-alpha: 1.2924 2
 Mo, K-beta: 0.946 0.246
 F(004): 256.0
 F(008): 256.41
 D, K, lambda: 6.15, 13 6.15, 13
 theta: 23.59 24.69
 Theta, lambda: 0.966, 0.366
 Theta: 0.966
 Correction method: 'MULTI-SCAN'
 Data completeness: 0.97 theta(max)= 07.49
 R(intensity)= 0.0531(193) wR (intensity)= 0.1292(249)
 R = 0.040 wR = 0.080



EPSRC National Crystallography Service
 Data Collection Summary

Summary report for Directory: diska/02sot082

Report generated Jul 09, 2002; 10:13:51

Unit cell

15124 reflections with 2.91° theta (2θ = 48°) were used for unit cell refinement

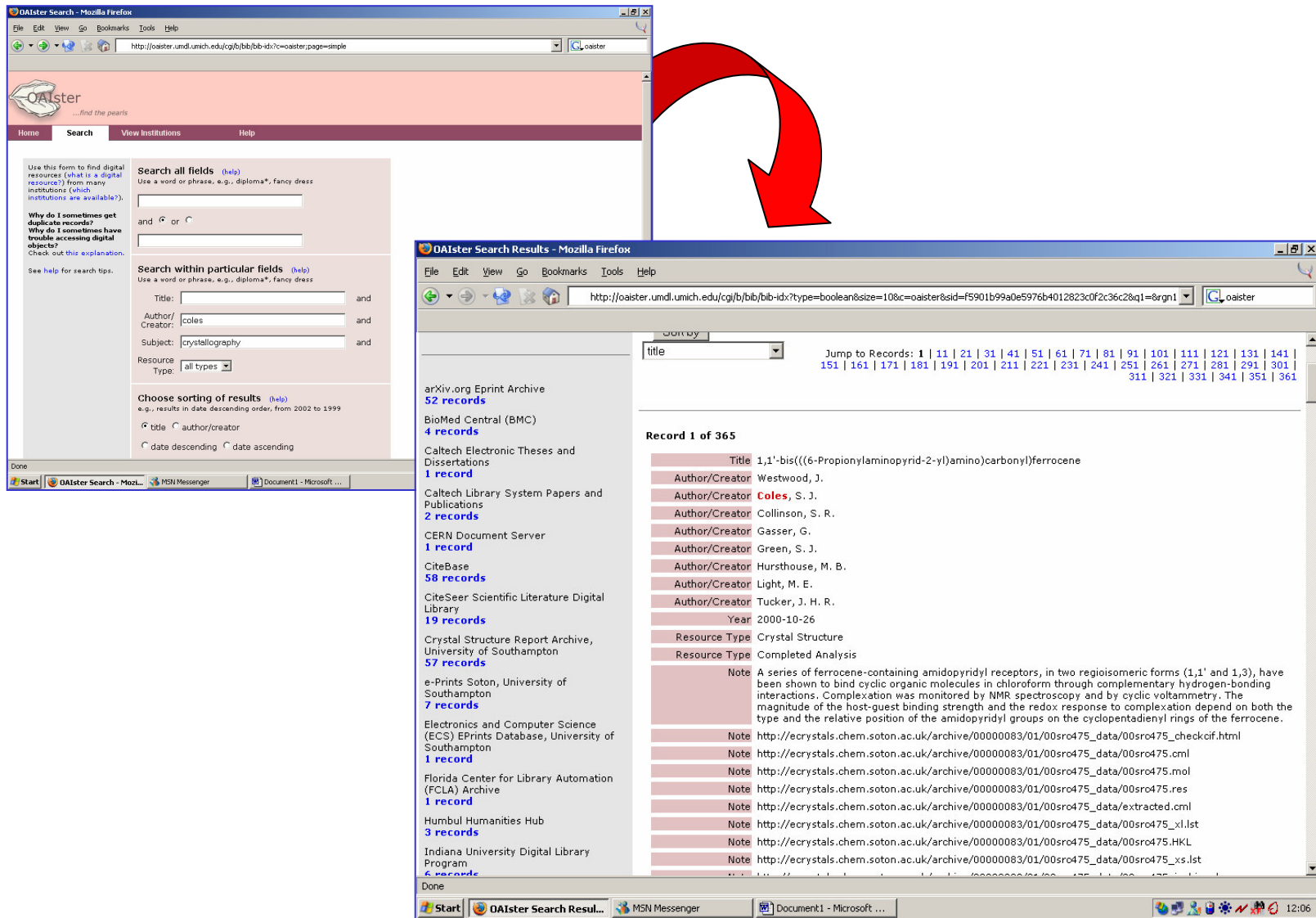
Symmetry used: P222

a (Angstroms): 9.3133 ± 0.0003
 b (Angstroms): 9.8424 ± 0.0003
 c (Angstroms): 15.4441 ± 0.0004
 alpha (°): 90.000
 beta (°): 90.000
 gamma (°): 90.000
 Volume (Å³): 1415.69 ± 0.07
 Molecular Weight: 0.743 ± 0.002

h	k	l	I	σ(I)
11	-1	1	1.08	9.10
12	-1	1	4.12	4.12
13	0	-2	10.79	9.22
14	0	-2	6.29	6.13
15	-2	12	4.18	4.18
16	2	-12	10.79	9.22
17	2	-12	6.29	6.13
18	-1	1	0.20	0.05
19	1	1	0.20	0.05
20	1	1	0.20	0.05
21	-4	5	40.42	5.15
22	4	-5	40.42	5.15
23	4	5	0.04	0.07
24	-4	-5	0.04	0.04
25	4	4	10.10	4.04
26	-4	-4	20.57	5.04
27	4	4	20.57	4.13
28	-4	-4	7.04	4.00
29	4	-4	7.04	4.00
30	4	4	0.04	0.07
31	-4	-4	0.04	0.06
32	4	4	0.04	0.07
33	-4	-4	0.04	0.07
34	4	4	0.04	0.07
35	-4	-4	0.04	0.07
36	4	4	0.04	0.07
37	-4	-4	0.04	0.07
38	4	4	0.04	0.07
39	-4	-4	0.04	0.07
40	4	4	0.04	0.07
41	-4	-4	0.04	0.07
42	4	4	0.04	0.07
43	-4	-4	0.04	0.07
44	4	4	0.04	0.07
45	-4	-4	0.04	0.07
46	4	4	0.04	0.07
47	-4	-4	0.04	0.07
48	4	4	0.04	0.07
49	-4	-4	0.04	0.07
50	4	4	0.04	0.07
51	-4	-4	0.04	0.07
52	4	4	0.04	0.07
53	-4	-4	0.04	0.07
54	4	4	0.04	0.07
55	-4	-4	0.04	0.07
56	4	4	0.04	0.07
57	-4	-4	0.04	0.07
58	4	4	0.04	0.07
59	-4	-4	0.04	0.07
60	4	4	0.04	0.07
61	-4	-4	0.04	0.07
62	4	4	0.04	0.07
63	-4	-4	0.04	0.07
64	4	4	0.04	0.07
65	-4	-4	0.04	0.07
66	4	4	0.04	0.07
67	-4	-4	0.04	0.07
68	4	4	0.04	0.07
69	-4	-4	0.04	0.07
70	4	4	0.04	0.07
71	-4	-4	0.04	0.07
72	4	4	0.04	0.07
73	-4	-4	0.04	0.07
74	4	4	0.04	0.07
75	-4	-4	0.04	0.07
76	4	4	0.04	0.07
77	-4	-4	0.04	0.07
78	4	4	0.04	0.07
79	-4	-4	0.04	0.07
80	4	4	0.04	0.07
81	-4	-4	0.04	0.07
82	4	4	0.04	0.07
83	-4	-4	0.04	0.07
84	4	4	0.04	0.07
85	-4	-4	0.04	0.07
86	4	4	0.04	0.07
87	-4	-4	0.04	0.07
88	4	4	0.04	0.07
89	-4	-4	0.04	0.07
90	4	4	0.04	0.07
91	-4	-4	0.04	0.07
92	4	4	0.04	0.07
93	-4	-4	0.04	0.07
94	4	4	0.04	0.07
95	-4	-4	0.04	0.07
96	4	4	0.04	0.07
97	-4	-4	0.04	0.07
98	4	4	0.04	0.07
99	-4	-4	0.04	0.07
100	4	4	0.04	0.07



Harvesting: OAIster



The image shows two screenshots of the OAIster search interface. The left screenshot shows the search form with the following fields and options:

- Search all fields:** A text input field for a search query.
- Search within particular fields:** Fields for Title, Author/Creator (with 'coles' entered), Subject (with 'crystallography' entered), and Resource Type (set to 'all types').
- Choose sorting of results:** Radio buttons for 'title', 'author/creator', 'date descending', and 'date ascending'.

The right screenshot shows the search results page for the query 'coles'. It displays a list of records from various institutions, including:

- arXiv.org Eprint Archive: 52 records
- BioMed Central (BMC): 4 records
- Caltech Electronic Theses and Dissertations: 1 record
- Caltech Library System Papers and Publications: 2 records
- CERN Document Server: 1 record
- CiteBase: 58 records
- CiteSeer Scientific Literature Digital Library: 19 records
- Crystal Structure Report Archive, University of Southampton: 57 records
- e-Prints Soton, University of Southampton: 7 records
- Electronics and Computer Science (ECS) EPrints Database, University of Southampton: 1 record
- Florida Center for Library Automation (FLA) Archive: 1 record
- Humbul Humanities Hub: 3 records
- Indiana University Digital Library Program: 6 records

The first record is expanded, showing the following details:

- Title:** 1,1'-bis(((6-Propionylaminopyrid-2-yl)amino)carbonyl)ferrocene
- Author/Creator:** Westwood, J.
- Author/Creator:** Coles, S. J.
- Author/Creator:** Collinson, S. R.
- Author/Creator:** Gasser, G.
- Author/Creator:** Green, S. J.
- Author/Creator:** Hursthouse, M. B.
- Author/Creator:** Light, M. E.
- Author/Creator:** Tucker, J. H. R.
- Year:** 2000-10-26
- Resource Type:** Crystal Structure
- Resource Type:** Completed Analysis
- Note:** A series of ferrocene-containing amidopyridyl receptors, in two regioisomeric forms (1,1' and 1,3), have been shown to bind cyclic organic molecules in chloroform through complementary hydrogen-bonding interactions. Complexation was monitored by NMR spectroscopy and by cyclic voltammetry. The magnitude of the host-guest binding strength and the redox response to complexation depend on both the type and the relative position of the amidopyridyl groups on the cyclopentadienyl rings of the ferrocene.
- Note:** http://ecrystals.chem.soton.ac.uk/archive/00000083/01/00sr0475_data/00sr0475_checkcif.html
- Note:** http://ecrystals.chem.soton.ac.uk/archive/00000083/01/00sr0475_data/00sr0475.cml
- Note:** http://ecrystals.chem.soton.ac.uk/archive/00000083/01/00sr0475_data/00sr0475.mol
- Note:** http://ecrystals.chem.soton.ac.uk/archive/00000083/01/00sr0475_data/00sr0475.res
- Note:** http://ecrystals.chem.soton.ac.uk/archive/00000083/01/00sr0475_data/extracted.cml
- Note:** http://ecrystals.chem.soton.ac.uk/archive/00000083/01/00sr0475_data/00sr0475_xl.lst
- Note:** http://ecrystals.chem.soton.ac.uk/archive/00000083/01/00sr0475_data/00sr0475.HKL
- Note:** http://ecrystals.chem.soton.ac.uk/archive/00000083/01/00sr0475_data/00sr0475_xs.lst



Aggregating: search & discover

eBank UK Demo

This is a prototype interface for the [eBank UK](#) JISC-funded project. It demonstrates an OAI-PMH aggregator service which cross searches a small sample of metadata records describing crystallography experiments (provided by the National Crystallography Service at the University of Southampton), and a small number of metadata records describing articles from the Crystallography literature (made available for use in this demo only by IUCr.) Links to the crystallography data sets and to the articles on line at the IUCr website are available in the search results.

Search for entries matching all the following:

Author =

CCDC Code =

IUPAC name =

Empirical Formula =

Compound Class =

General keywords =

Date released =

OR published in the last

Search within: Data Reports Publications e.g. journal articles

eBank UK Demo

This is a prototype interface for the [eBank UK](#) JISC-funded project. It demonstrates an OAI-PMH aggregator service which cross searches a small sample of metadata records describing crystallography experiments (provided by the National Crystallography Service at the University of Southampton), and a small number of metadata records describing articles from the Crystallography literature (made available for use in this demo only by IUCr.) Links to the crystallography data sets and to the articles on line at the IUCr website are available in the search results.

Search for entries matching all the following:

Author =

CCDC Code =

IUPAC name =

Empirical Formula =

Compound Class =

General keywords =

Date released =

OR published in the last

Search within: Data Reports Publications e.g. journal articles



Linking data to publications

eBank UK Demo

Crystal Structure Data Reports

[Crystal Structure Report of 2-\(N-Ferrocenylmethylcarbamoyl\)-5-\(N-phenylcarbamoyl\)-3,4-diphenyl pyrrole](#)
Creator(s): Hursthouse, Michael B., Light, Mark E., Coles, Simon J., Horton, Peter N., Gale, Phil A., Denuault, G., Warriner, C. N.
Date released: 23/05/2004
Empirical Formula: C35H29F6N3O2
IUPAC name: 2-(N-Ferrocenylmethylcarbamoyl)-5-(N-phenylcarbamoyl)-3,4-diphenyl pyrrole
CCDC code: XU2SIU
Compound Class: Organic
General keywords: Supramolecular Chemistry
Related article: [?A URI citation?](#)
Available Datasets
CIF file
processing Dataset
refinement Dataset
solution Dataset

[Crystal Structure Report of 2-\(N-Ferrocenylcarbamoyl\)-5-\(methoxycarbonyl\)-3,4-diphenylpyrrole](#)
Creator(s): Hursthouse, Michael B., Coles, Simon J., Light, Mark E., Horton, Peter N., Gale, Phil A., Denuault, G., Warriner, C. N.
Date released: 23/05/2004
Empirical Formula: C29H24F6N3O3
IUPAC name: 2-(N-Ferrocenylcarbamoyl)-5-(methoxycarbonyl)-3,4-diphenylpyrrole
CCDC code: XU2SOA
Compound Class: Organometallic
General keywords: Supramolecular Chemistry
Related article: [?A URI citation?](#)
Available Datasets
CIF file
processing Dataset
refinement Dataset
solution Dataset

Publications

A supramolecular assembly: aquatris(pentafluorophenyl)borane as its mixed dimethyl sulfone and water solvate, (H₂O)₃(C₆F₅)₃Me₂S₂O₂H₂O

The title compound, C₁₈H₂BF₁₅O₄H₂O₂, obtained by crystallization of a product formed from a reaction mixture containing 5(C₆F₅)₃ and Me₂S₂O₂ (and H₂O) in hexane, was characterized in the solid state as a supramolecular assembly containing water adducts of tris(pentafluorophenyl)borane, (H₂O)₃(C₆F₅)₃, linked together by a network of hydrogen bonds involving one additional H₂O and one additional Me₂S₂O₂ molecule per adduct molecule.
Creator(s): Coles, Simon J., Hursthouse, Michael B., Beckett, Michael A., Dutton, Michael Acta Crystallogr E Struct Rep Online Vol 59 Issue Pt 9 pp. o1354 - o1356
DOI:
Download from: <http://scripts.iucr.org/cgi-bin/getarticleid?issn=1600-5368&volume=59&page=1354&details=yes>

Structural investigations of phosphorus–nitrogen compounds. 5. Relationships between molecular parameters of 2,2-diphenyl-4,6-cis-oxytetra(ethyleneoxy)-4,6-R₂-cyclophosphazatrienes (R = Cl, OCH₂CF₃, OPh, OMe, NHPH, NHtBu) and substituent basicity constants

The syntheses and crystal structures of six new cis-ansa derivatives N3P3Ph2[O(CH₂CH₂O)]₄ R₂ (R = Cl, OCH₂CF₃, OPh, OMe, NHPH, NHtBu) are reported and the observed relationship between molecular parameters of the N3P3 ring and substituent basicity constants is discussed.
Creator(s): Besli, S., Coles, S. J., Hursthouse, M. B., Kilic, A., Mayer, T. A., Shaw, R. A. Acta Crystallogr B Vol 58 Pt 6 pp. 1067 - 1073
DOI: 10.1107/S0108768102018608
Download from: <http://scripts.iucr.org/cgi-bin/getarticleid?issn=0108-7681&volume=58&page=1067&details=yes>
Related dataset: <http://ecrystals.chem.soton.ac.uk/archive/00000062/>

5 α -Cholestanol

The title compound, C₂₇H₄₈, is a steroid derivative composed of a saturated-carbon fused-ring framework with two methyl substituents and an alkyl side chain.
Creator(s): Coles, S. J., Hursthouse, M. B., Frampton, C. S. Acta Crystallogr E Struct Rep Online Vol 58 Issue Pt 4 pp. o445 - o446
DOI: 10.1107/S1600536802004786
Download from: <http://scripts.iucr.org/cgi-bin/getarticleid?issn=1600-5368&volume=58&page=445&details=yes>
Related dataset: <http://ecrystals.chem.soton.ac.uk/archive/00000051/>

Ethyl (2S⁺)-2-[(2R⁺,2R⁺,5S⁺)-2',5'-dimethyl-5'-oxoperhydro-[2,2']bifuranyl-5-yl]-2-hydroxyethanolate

The framework of K₂Zn(H₂P₂O₇)₂·2H₂O contains acid diphosphate–metallate layers linked by KO interactions and weak hydrogen bonds. Zn²⁺ cations are coordinated octahedrally by O atoms from two bidentate [H₂P₂O₇]²⁻ anions and two water molecules.

Benzene 1,2dicarboxylic acid

Simon J. Coles, Michael B. Hursthouse, Claire L. Taylor and Peter N. Horton.
University of Southampton
C₈H₆O₄

ICN Code: INChI=1.12Beta/C8H6O4c9-7|10|5-3-1-2-4-6|58(11)2h1-4h|(H,9,10)(H,11,12) (google for ich)
Compound Class: Organic
Keywords: Phthalic acid
Creation Date: 15 February 2005
Deposited By: Dr Simon J. Coles
Deposited On: 21 February 2005

Data collection parameters

Chemical formula	C ₈ H ₆ O ₄
Crystallisation Solvent	
Crystal morphology	Prism
Crystal system	monoclinic
Space group symbol	C2/c
Cell length a	5.0018(10)
Cell length b	14.214(3)
Cell length c	9.5198(19)
Cell angle alpha	90.00
Cell angle beta	94.33(3)
Cell angle gamma	90.00
Data collection temperature	120(2)

Available Files

Final Result	
05mh1006.cmi	3k
05mh100605mh1006.cif	9k
05mh100605mh1006_checkcif.htm	7k
05mh1006_ichsi.cmi	1k

Refinement

05mh100605mh1006.res	3k
05mh100605mh1006_xlst	21k

research papers

Acta Crystallographica Section B
Structural
Science
ISSN 0108-7681

Structural investigations of phosphorus–nitrogen compounds. 6. Relationships between molecular parameters in per-X-substituted bridged spermine derivatives and basicity constants $\Sigma\alpha R$ of substituents

Simon J. Coles,^a David B. Davies,^b Michael B. Hursthouse,^a Adem Kilic,^c Thomas A. Mayer,^a Robert A. Shaw^b and Gonul Yenilmez Girci^c

^aSchool of Chemistry, University of Southampton, Highfield, Southampton SO17 1BE, England, ^bSchool of Biological and Chemical Sciences, Birkbeck College, University of London, Gordon House, 29 Gordon Square, London WC1H 0PP, England, and ^cDepartment of Chemistry, Gazi Institute of Technology, Gebze, Turkey

Correspondence email: s.j.coles@soton.ac.uk

Received 8 July 2004
Accepted 13 October 2004

A systematic study is reported of the products of the nucleophilic substitution reactions of the spermine-bridged cyclophosphazene, [N₃P₃X(NHCH₂CH₂CH₂NHCH₂CH₂)]₄ [where X = Cl (2a)], to give a number of new structures [(2b)–(2g)] in which X = OPh, [spiro-O(CH₂)₃O]₂, Ph, NHPH, NC₄H₉ and NHtBu^d, respectively. A comparison has been made between the sum of the substituent basicity constants, $\Sigma\alpha R$, obtained in nitrobenzene solution, and ten molecular parameters of the N₃P₃ ring (the internal bond angles α , β , γ , δ and θ , and the P–N bond lengths a , b , c , d and e) as well as the difference between the bond lengths a and b , $\Delta(P-N)$. It is found that the systematic change in molecular parameters of compounds (2b)–(2g) is in line with changes in αR values, indicating the similarity in relative electron-releasing capacity of substituents X in the solid state and in solution. It is also found that the effect on molecular parameters of (2b)–(2g) with two X substituents in P₃ groups is greater than that for one X substituent in P(O)X groups in an analogous series of compounds observed previously [Besli *et al.* (2002), *Acta Cryst. B* **58**, 1067–1073].



Embedding in a science portal for student learners



PSIgate
Physical Sciences Information Gateway

SEARCH SUBJECTS

- Astronomy
- Chemistry
- Earth sciences
- Materials science
- Physics
- History & policy

FEATURES

- PSIgate home
- Search PSIgate
- Subject Headings A-Z
- About PSIgate
- Site map
- New additions
- Feedback form
- Suggest site
- News services
- Reference
- Spotlight
- Science Timelines
- Hot Topics
- Science Data
- Courses
- Jobs
- Learning & Teaching

PSIgate Home > eBank

Your search returned 28 data reports and 4 publications. Viewing 1 to 10

[Next](#)
[New search](#)

Crystal Structure Data Reports

[Crystal Structure Report of 2-\(N-Ferrocenylmethylcarbamoyl\)-5-\(N-phenylcarbamoyl\)-3,4-diphenyl pyrrole](#)

Creator(s): Hursthouse, Michael B., Light, Mark E., Coles, Simon J., Horton, Peter N., Gale, Phil A., Denuault, G., Warriner, C. N.

Date released: 23/05/2004

Empirical Formula: C35H29FeN3O2

IUPAC name: 2-(N-Ferrocenylmethylcarbamoyl)-5-(N-phenylcarbamoyl)-3,4-diphenyl pyrrole

Compound Class: Organic

General keywords: Supramolecular Chemistry

Related article: [?A URI citation?](#)

Copyright © PSIgate. All rights reserved. PSIgate is a service of the [Resource Discovery Network \(RDN\)](#). feedback@psigate.ac.uk



Ontologies for discovery in an inter-disciplinary world

A	List of Controlled Keywords
ABCDEF	ABCDEFGHIJKLMN
OPQRST	UVWXYZ
A	
I	ab-initio calculations +
J	ab-initio periodical and cluster calculations +
K	ab-initio powder structure determination +
L	ab-initio structure determination +
M	ablation +
N	absolute chirality +
O	absolute configuration +
P	absolute configuration determination +
Q	absolute configuration organic compounds +
R	absolute polarity +
S	absolute structure +
T	absolute structure determination +
U	absolute structure factors +
V	absorption +
W	absorption correction +
X	absorption edge +
Y	absorption spectroscopy +
Z	absorption spectroscopy experimental +
	absorption spectroscopy theoretical +
	academic management +
	accuracy +
	accurate data collection +
	accurate data processing +

- Transform the 'list' into an 'ontology'
- Embed ontology into the deposition process
- Aggregators use keywords for linking with the broader literature
- Researchers use keyword ontology in search and discovery services

Persistent identifiers for data citation

- eBank use cases: depositor, author, service provider, reader, publisher, ?
- Schemes: DOI, Handle, ARK, PURL
- Global identification: express as http URIs
- Added value services: CrossRef, resolution service, integration (Globus), look-up service, ?
- Degree of trust or persistence
- Costs
- Future potential: political, ?
- Domain identifiers: International Chemical Identifier (InChI) codes



Publication & citation of scientific primary data project

- National Library for Science & Technology (TIB), University of Hanover, Germany
- STD-DOI Project *<http://www.std-doi.de>*
- DOI registry for datasets
- Data requirements: quality control, long-term curation, use DOI resolver
- Data publication agents: World Data Center Climate, GeoForschungsZentrum Potsdam
- Exemplar data citation:
 - *Kamm, H; Machon, L; Donner, S (2004): Gas chromatography (KTB Field Lab), GFZ Potsdam.*
doi:10.1594/GFZ/ICDP/KTB/ktb-geoch-gaschr-p



Foundations of
Crystallography
Online



current issue
back issues
sample issue

reader services
author services
article submission

Structural
Science
Online



current issue
back issues
sample issue

reader services
author services
article submission

Crystal Structure
Communications
Online



current issue
back issues
sample issue

reader services
author services
article submission

Biological
Crystallography
Online



current issue
back issues
sample issue

reader services
author services
article submission

Publishers
seal of
approval



Integration into crystallographic publishing practices

University of Southampton Crystal Structure Report Archive

Home
About
Browse
Search
Register
User Area
Help

Welcome to Southampton EBank

Southampton eCrystals is the archive for Crystal Structures generated by the Southampton Chemical Crystallography Group and the EPSRC UK National Crystallography Service.

The information contained within each entry of this archive is all the fundamental and derived data resulting from a single crystal X-ray structure determination, but excluding the raw images. The results have not been externally refereed, but the information supplied should enable any reader to check the reliability and validity directly, since all the files provided are freely available for download. Should any error be detected, we would appreciate receiving suitable comments, and we will make any necessary amendments, and include a note to that effect. Any reader wishing to have access to the raw images is welcome to contact us, and we will make arrangements for these to be made available.

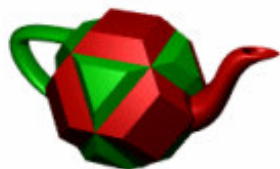


Integration into chemistry research workflows

R4L Repository for the Laboratory

- R4L Repository for the Laboratory Project (JISC-funded) automated data capture from instrumentation, registration of results
- SMART TEA electronic Laboratory notebook + annotations

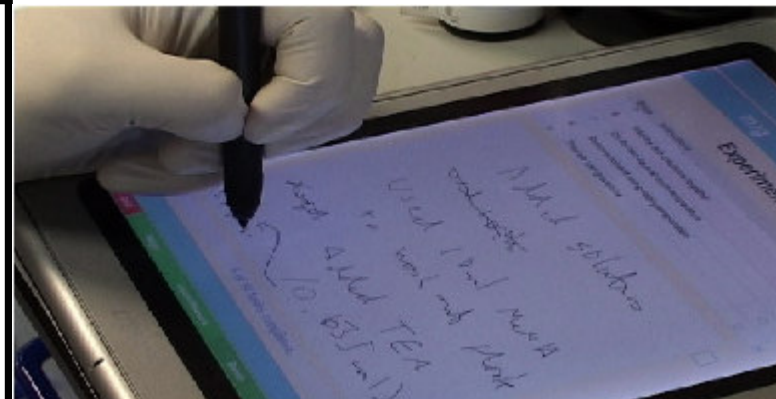
the Smart Tea Project



"I can go anywhere and its, like, this is me and my data. Its all there, bang."

- Chris,
a real chemist, on using Smart Tea
instead of a paper lab book.

Smart Tea is about improving the information environment for chemists doing chemistry - within and beyond the lab. Smart Tea is about supporting chemists in the preparation, execution, analysis and dissemination of their experimental work.



the myTea project

- Related sub-domains of chemistry: SPECTRa Project (JISC-funded)
- Research assessment (RAE) process?



Integration into the curriculum and e-Learning workflows

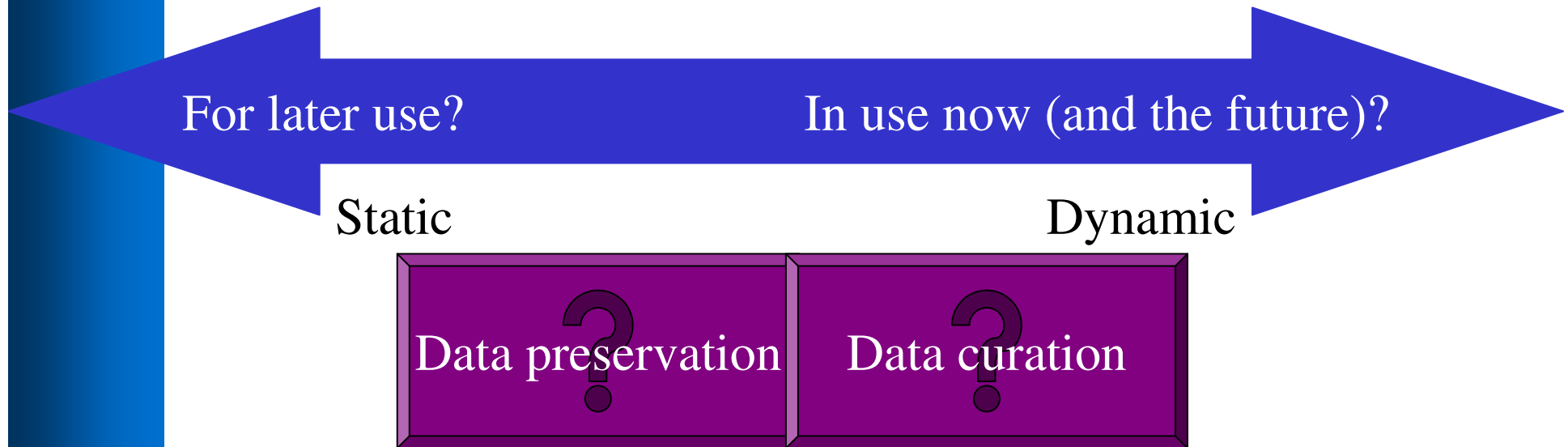
- MChem course
- Assess role in Undergraduate Chemical Informatics courses
- Pedagogic evaluation



3. Looking to the longer term: digital curation & preservation



Repositories and digital curation



“maintaining and adding value to a trusted body of digital information for current and future use”

Assuring long term access to the research record

- Trusted digital repositories
 - Audit Checklist for Certification Draft Report
 - Research Libraries Group, August 2005
 - RLG-NARA Taskforce
 - Defined criteria under 4 categories
 - Organisation
 - Functions, processes & procedures
 - Designated community & usability
 - Technologies & technical infrastructure
- UK Digital Curation Centre *<http://www.dcc.ac.uk>*
 - 1st International DCC Conference presentations available
 - PV2005 Royal Society Edinburgh November 21-23 Nov



Thank you.
Questions?.....

More information: UKOLN <http://www.ukoln.ac.uk/>



