

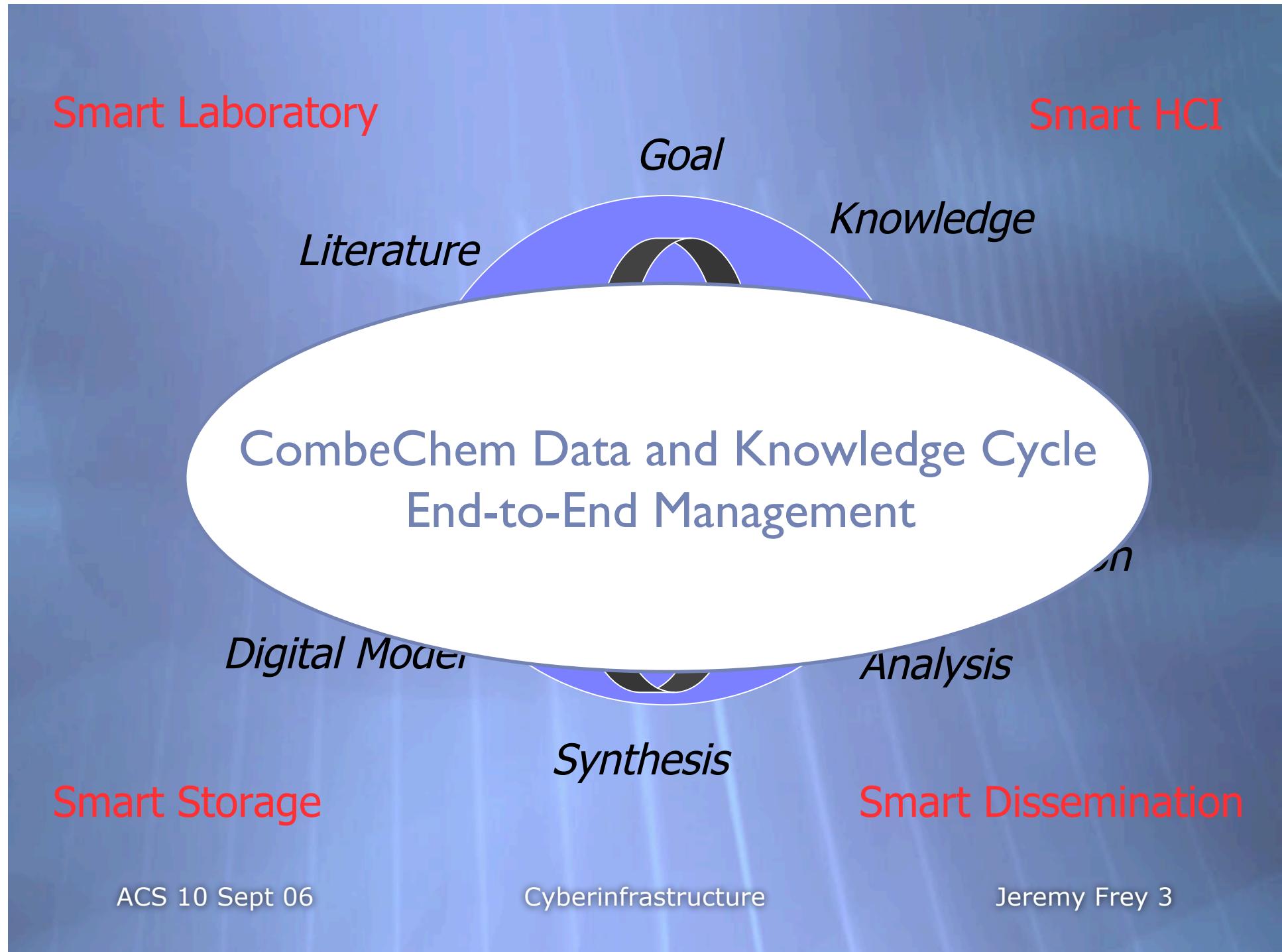
The “CombeChem” Project

Semantic Support for the Chemical Information
Life cycle

Jeremy Frey
School of Chemistry
University of Southampton

The CombeChem Project

- ★ End to End linking of data and information
 - ★ Laboratory to publication and back again
 - ★ Very long data chains can be involved e.g. from a chemistry lab to mouse genetic expression
- ★ The exponential world of combinatorial synthesis and high throughput analysis meets the exponentially growing power of computing
 - ★ “Automation, Semantics & the Grid”



e-Science

- ★ 'e-Science is about global collaboration in key areas of science, and the next generation of infrastructure that will enable it.'

John Taylor, DG of UK OST

- ★ '[The Grid] intends to make access to computing power, scientific data repositories and experimental facilities as easy as the Web makes access to information.'

Tony Blair, 2002

- ★ What is the web?

The concept of Publication@Source

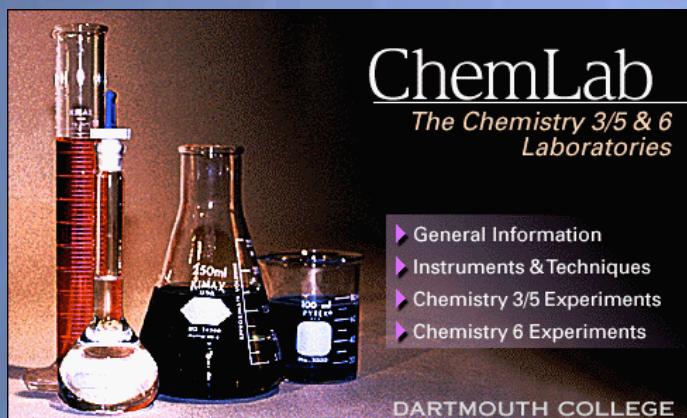
- ★ Trace all the way back from publication to the original data - provenance
- ★ The data is the key - DataGrid
- ★ Start as you mean to go on - ELNs are a necessity

necessary if a calculation or discussion is changed; the section to be deleted is simply removed by drawing a neat "x" through it.

In view of the fact that a notebook is a primary record, data are not copied into it from other sources (such as this manual or a lab partner's notebook, in a joint experiment) without clear acknowledgment of the source. Observations are never collected on note pads, filter paper, or other temporary paper for later transfer into a notebook. If you are caught using the "scrap of paper" technique, your improperly recorded data may be confiscated by your TA or instructor at any time. It is important to develop a standard approach to using a notebook routinely as the primary receptacle of observations.

Each week at the beginning of lab lecture, you will turn in your prelab problems from the manual for grading. Problems not turned in at the beginning of lab lecture will be

Observations are never collected on note pads, filter paper or other temporary paper for later transfer into a notebook



If you are caught using the "scrap of paper" technique, your improperly recorded data may be confiscated by your TA

Cyberinfrastructure

Lab books are a big block to publication@source: if it's not digital, it is more difficult to share



This is where it all starts: The Lab & The Lab Book

Need a usable digital lab book. Design by analogy to help Chemists and Computer Scientists work together.

ACS 10 Sept 06

Cyberinfrastructure



COSHH

leverage off things we already have to do

COSHH ASSESSMENT FORM				Record No.
SUBSTANCE NAME	PHYSICAL FORM	QUANTITY	NATURE OF HAZARD	
Water	liquid	1000ml	None	
Dextrose	soln	<20g	possible irritation to eyes and skin	
Caffeine	Solid (tea)	<1g	Harmful if swallowed, induce vomiting.	
Milk	liquid	<100ml	No particular hazards	
NATURE OF PROCESS				
Liquid extraction of caffeine, followed by combination with dextrose to produce a sweet drink				
Is there a less hazardous substance? No If so, why not use it?				
CONTROL MEASURES REQUIRED <i>(Local exhaust ventilation, personal protection, etc.)</i> No specific measure required				

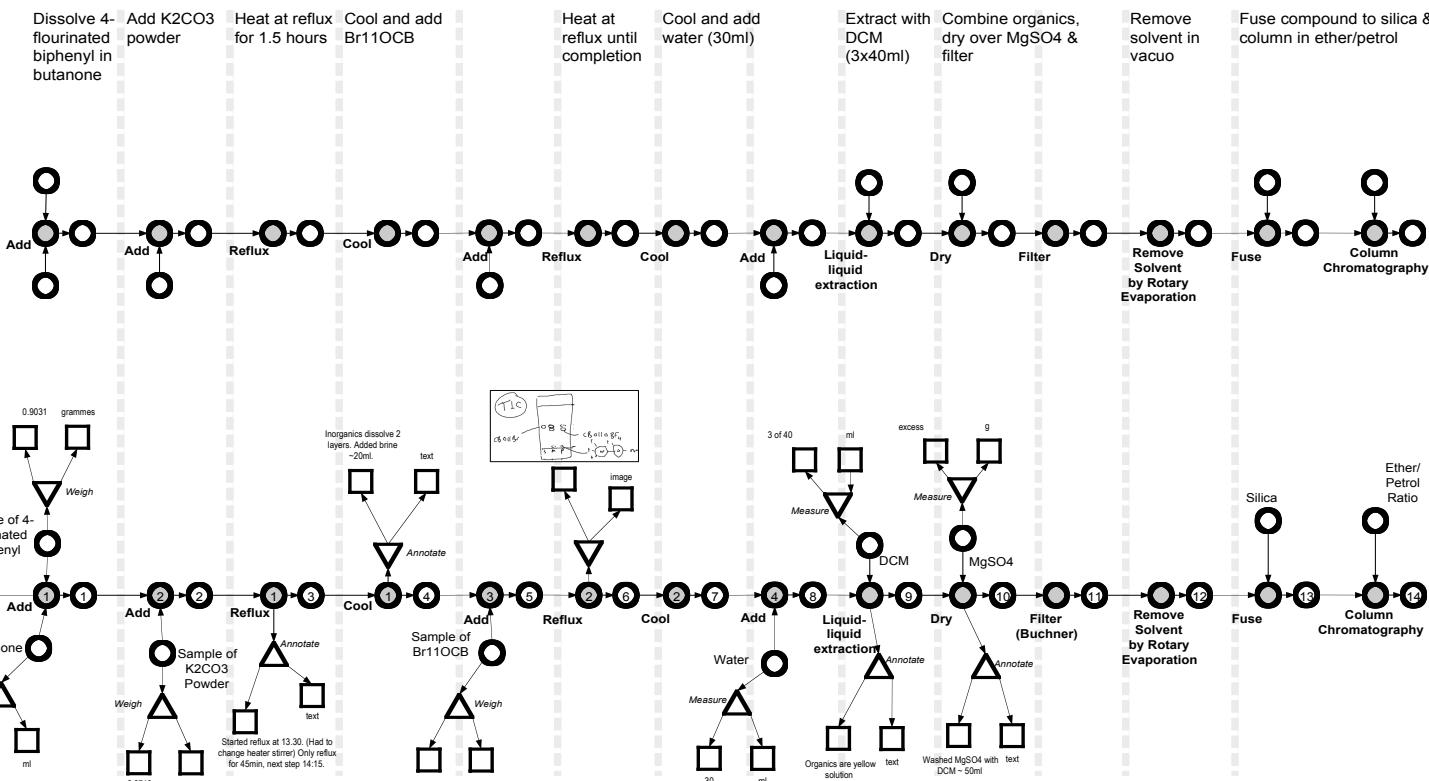
Perhaps computational chemists should be made to do the same!
Statistical modelling and design requires planning to be most effective

Process Record

To
List

PLAN

Ingredient List	
Fluorinated biphenyl	0.9 g
Br11OCB	1.59 g
Potassium Carbonate	2.07 g
Butanone	40 ml



Key	
Process	○
Input	○
Literal	□
Observation	▽

Observation Types	
weight - grammes	
measure - ml, drops	
annotate - text	
temperature - K, °C	

Future Questions	
Whether to have many subclasses of processes or fewer with annotations	
How to depict destructive processes	
How to depict taking lots of samples	
What is the observation/process boundary? e.g. MRI scan	

Combechem
30 January 2004
gvh, hrm, gms

ACS

Ingredient List

Fluorinated biphenyl	0.9 g
Br11OCB	1.59 g
Potassium Carbonate	2.07 g
Butanone	40 ml

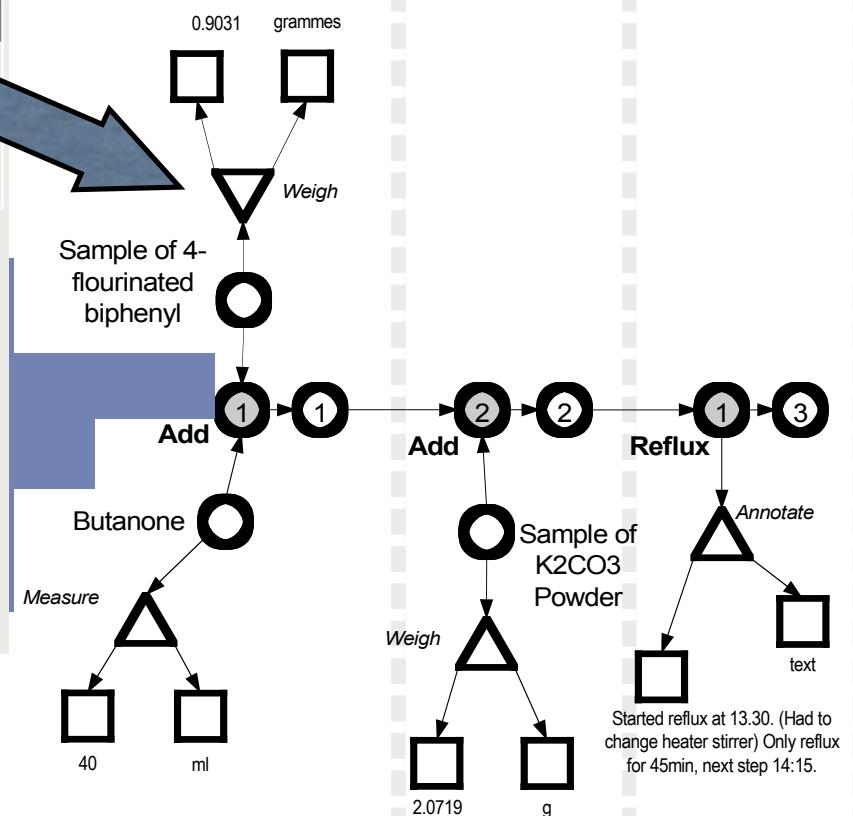
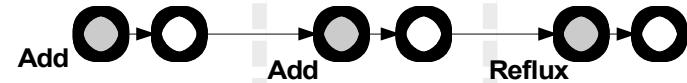
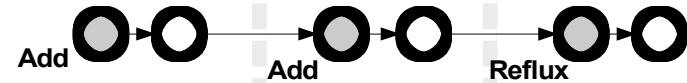
Name	Planned	Actual
Fluorinated biphenyl	0.9000 g	0.9031 g
Br11OCB	1.5900 g	1.5918 g
Potassium Carbonate	2.0700 g	2.0719 g
Butanone	40.0 ml	40.0 ml

7	8	9
4	5	6
1	2	3
0	.	

Enter**Del**

ACS 10 Sept 06

Dissolve 4-fluorinated biphenyl in butanone
Add K₂CO₃ powder
Heat at reflux for 1.5 hours

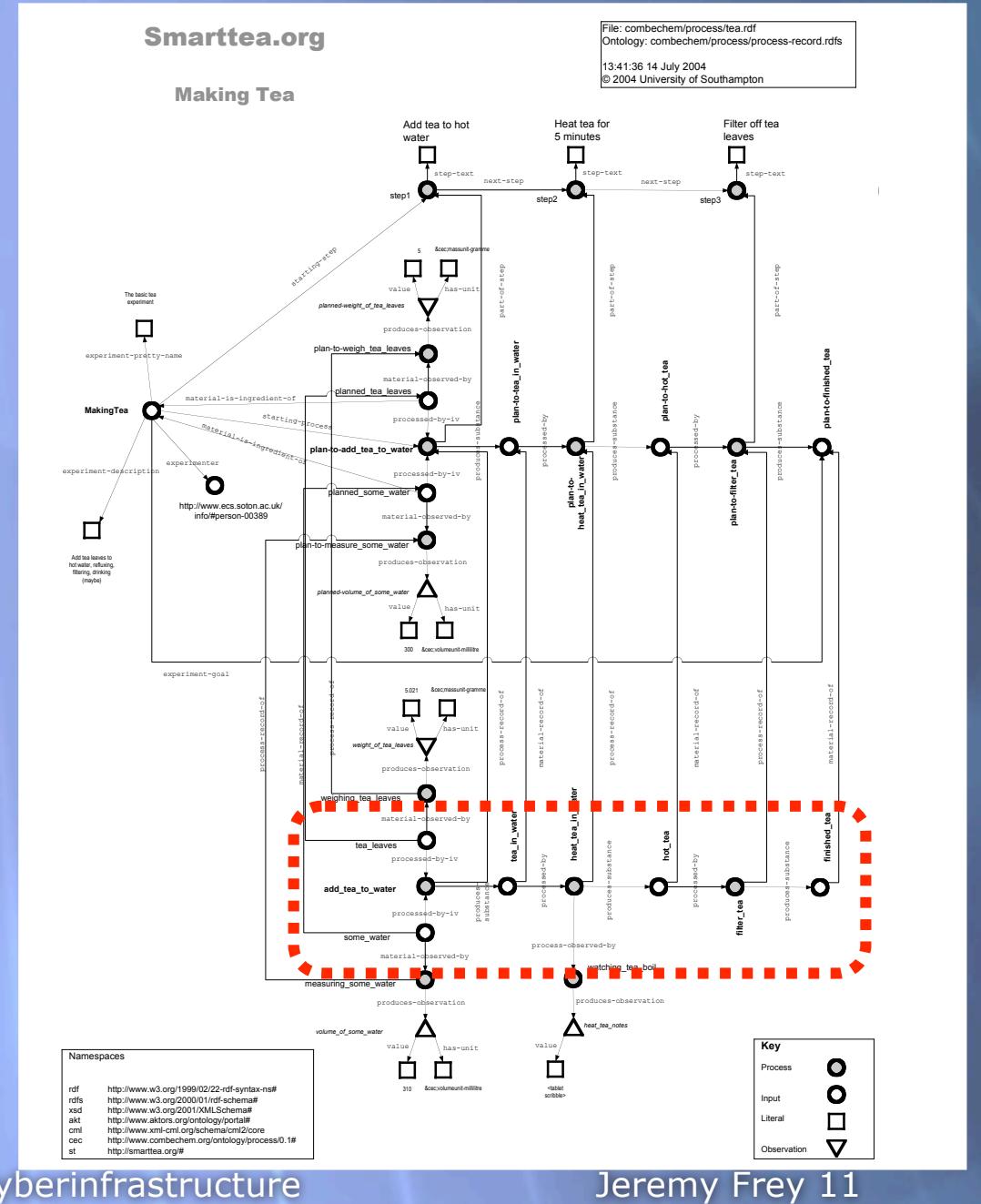


getRecord()

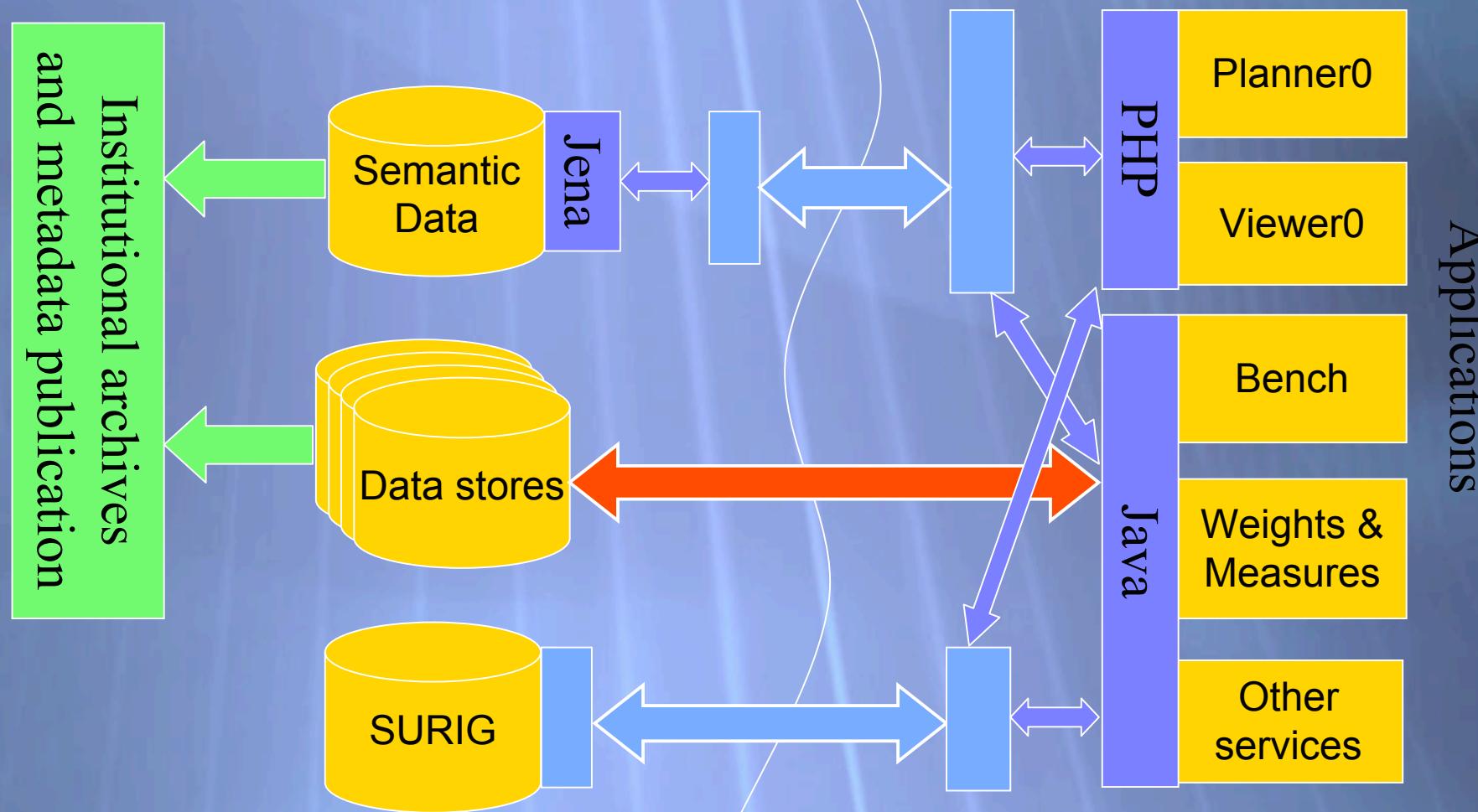
There is a potential containment problem in pulling back partial RDF graphs from the triple store.

Solved by using multiple triple stores but boundaries are a major issue for the future.

ACS 10 Sept 06



Architecture

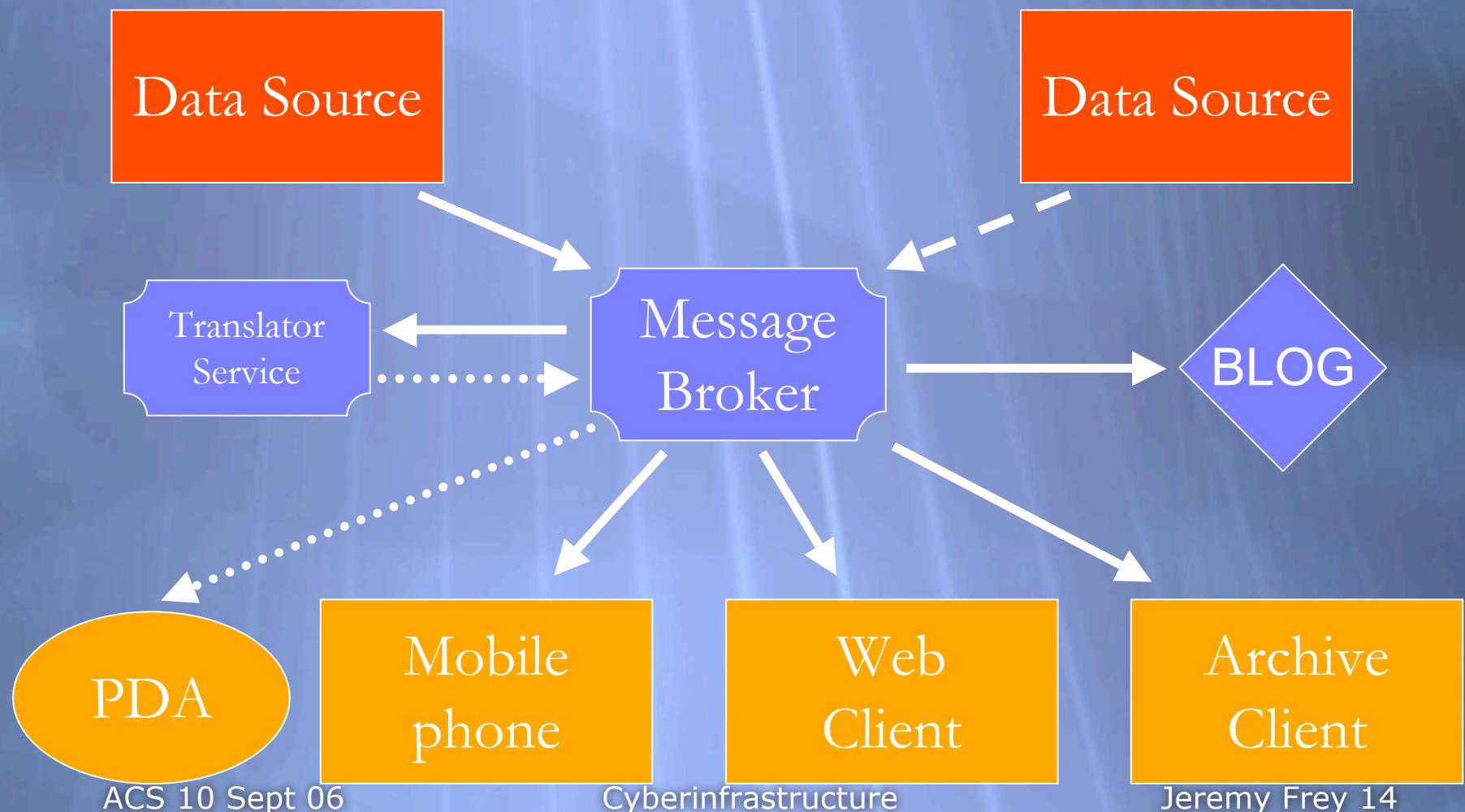


The Laboratory

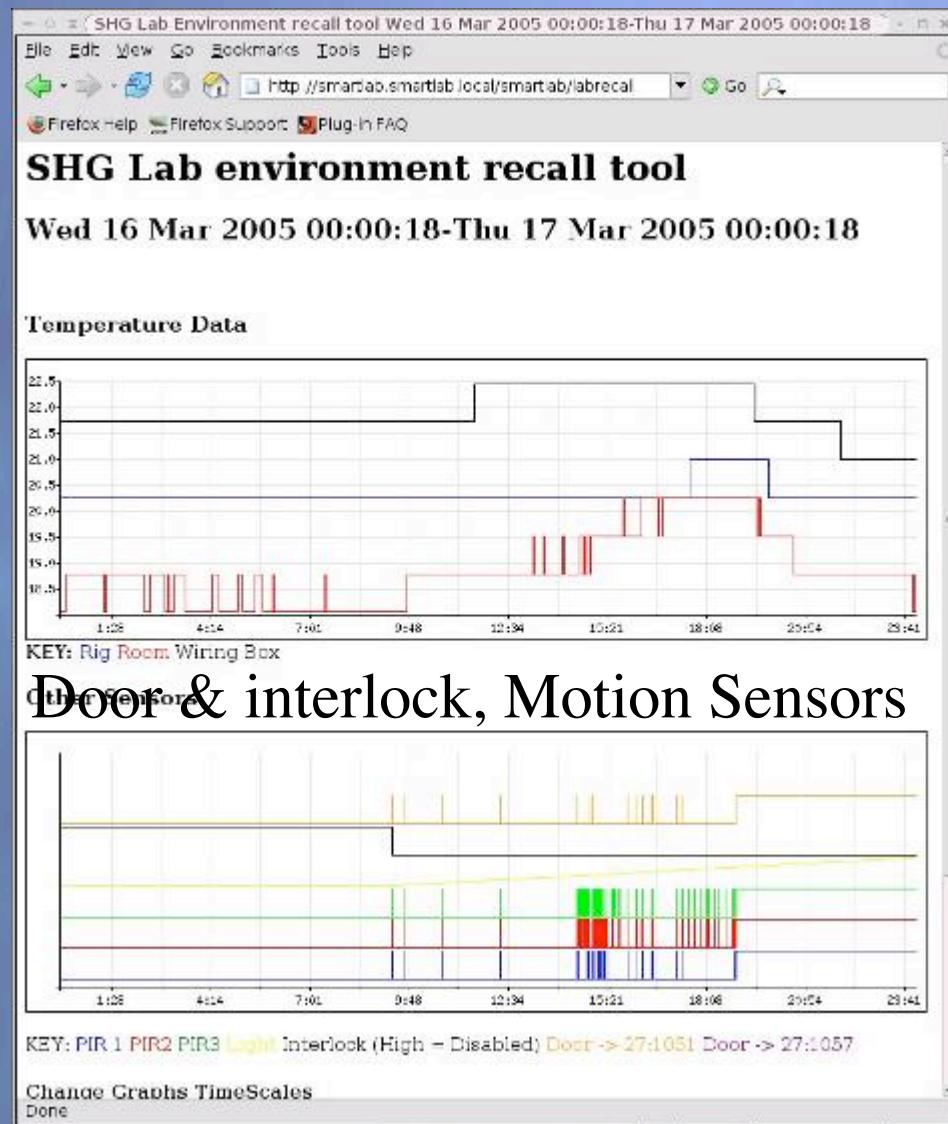
- ★ Capture information from places you would not want to put your eyes
- ★ Capture environmental data automatically
- ★ Capture people and movements
- ★ Provide this information in real time as well as for the laboratory record



Pub-Sub systems provide the flexible & extensible approach to distribution

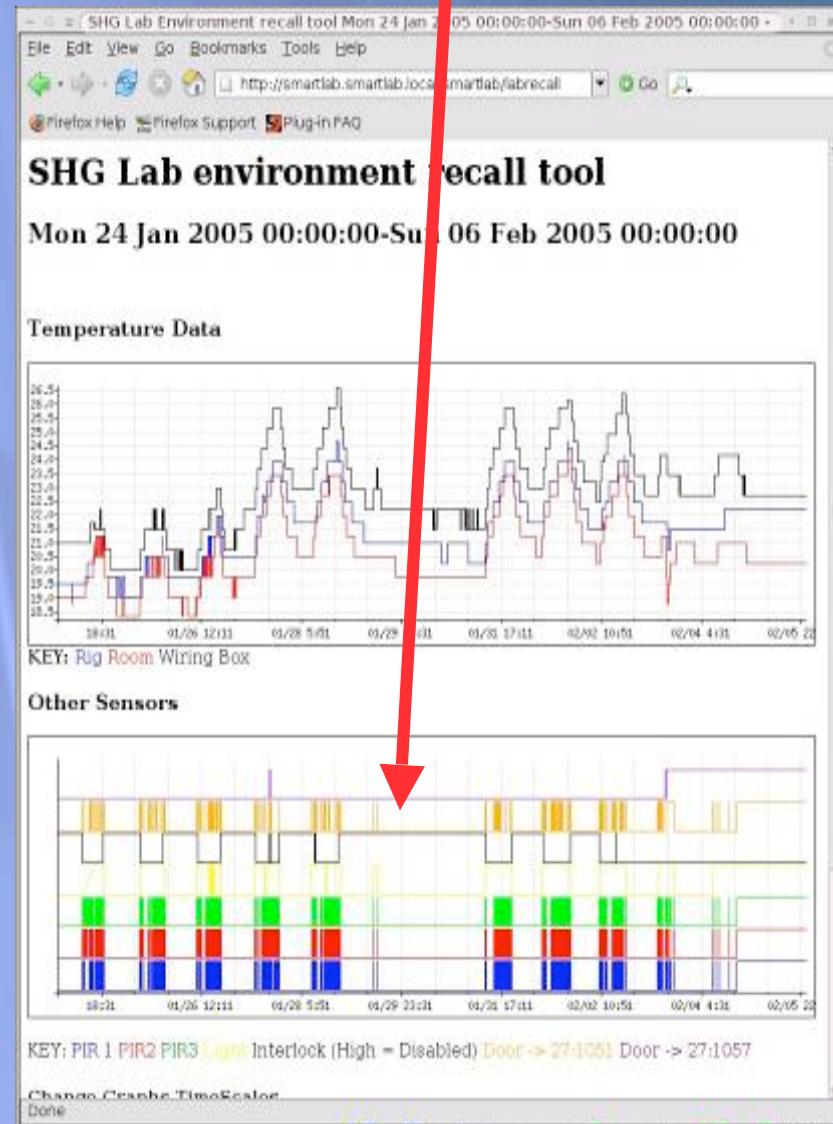


Temperature – room, laser



ACS 10 Sept 06

Air Conditioning failed



Jeremy Frey 15

Cyberinfrastructure

File Edit View Favorites Tools Help

Back → X

Address <http://news.bbc.co.uk/1/hi/sci/tech/4233757.stm>

bbc.co.uk

Low Graphics version | Change edition



News Front Page

World

UK

England

Northern Ireland

Scotland

Wales

Business

Politics

Health

Education

Science/Nature

Technology

Entertainment

Have Your Say

Magazine

In Pictures

Week at a Glance

Country Profiles

In Depth



File Edit View Favorites Tools Help

Back → X

Address <http://news.bbc.co.uk/1/hi/sci/tech/4233757.stm>

experiment in more detail.

PDAs everywhere

Members of Dr Frey's non-linear laser spectroscopy group, working as part of the Combechem project, successfully trying out the system at conferences and the university bar.

The next step is to evolve the system so that at a button on a mobile, scientists will be able to remote control the conditions in the lab, like turning down the tem

It could also realise the promise of controlling home environments from phones easily and seamlessly.

But there are many other potential applications for the monitoring jobs done by the power industries, healthcare professions and other labs.

IBM won the Royal Academy of Engineering's MacRobert prize which rewards technological and engineering innovation for the program in June last year.

Used by top global banks, the WebSphere MQ family is a decade old.



Chemists enjoy a drink at the bar while keeping an eye on the lab

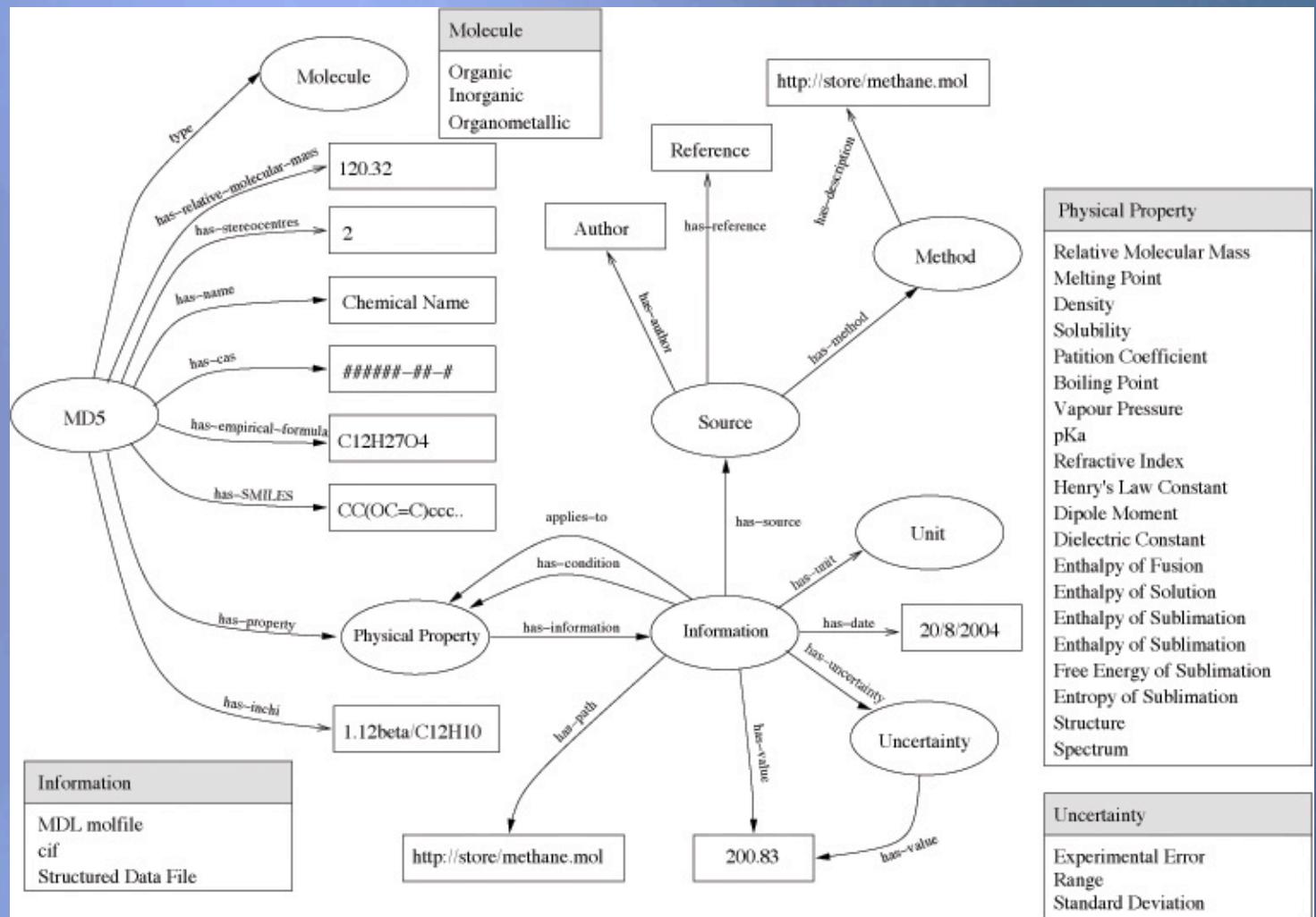


Databases - Our experience

- ★ What do you do when the actual users keep changing their mind?
- ★ Is a traditional relational database suitable?
- ★ Danger of re-enforcing scientific bias against relational database for laboratory data.
- ★ RDF & Triple stores were again the solution

RDF/RDFS

High level Schema for chemical properties



Property in RDF

- ★ <c:OrganicMolecule rdf:about="file:///storage/ba8efc2ce0edada69d63b02d1b8630c6.rdf">
- ★ <c:has-inchi>1.12Beta/C12H13NO2/c1-2-15-8-9-5-6-11(14)12-10(9)4-3-7-13-12/h1H3,2H2,3-7H,8H2,14H</c:has-inchi>
- ★ <c:has-cas>22049-19-0</c:has-cas>
- ★ <c:has-empirical-formula>C12H13NO2</c:has-empirical-formula>
- ★ <c:has-stereocentres>0</c:has-stereocentres>
- ★ <c:has-property>
- ★ <c:MeltingPoint>
- ★ <c:has-information>
- ★ <c:Information>
- ★ <c:has-value>150
- ★ </c:has-value>
- ★ <c:has-uncertainty>
- ★ <c:Range>
- ★ <c:has-value>16
- ★ </c:has-value>
- ★ </c:Range>
- ★ </c:has-uncertainty>
- ★ <c:Information>
- ★ </c:has-information>
- ★ </c:MeltingPoint>
- ★ <c:has-property>
- ★ ACS 10 Sept 06
- ★ </c:OrganicMolecule>

- ★ <rdfs:Class rdf:about="&c;OrganicMolecule">
 - ★ <rdfs:label>Organic Molecule</rdfs:label>
 - ★ <rdfs:subClassOf rdf:resource="&c;Molecule" />
 - ★ </rdfs:Class>
- ★ <rdfs:Class rdf:about="&c;PhysicalProperty">
 - ★ <rdfs:label>Property</rdfs:label>
 - ★ </rdfs:Class>
- ★ <rdfs:Class rdf:about="&c;PartitionCoefficient">
 - ★ <rdfs:label>Partition Coefficient</rdfs:label>
 - ★ <rdfs:subClassOf rdf:resource="&c;PhysicalProperty" />
 - ★ <rdfs:description>Ratio of substance dissolved in octan-1-ol and water</rdfs:description>
 - ★ </rdfs:Class>

File Edit View Go Bookmarks Tools Window Help

file:///rob/prog/triangle/test2/nitrostyrene-page.htm

File Edit View Go Bookmarks Tools Window Help

file:///rob/prog/triangle/test2/nitrostyrene-page.htm Search

TRIANGLE

CAS INCHI NAME SMILES EMPIR

1202-32-0

Search the DB Canned Example

query number = [10]

• MOLECULE

type	Resource
type	Molecule
has-SMILES	<chem>CCc1cc(C=C(N)O)O</chem>
has-name	B-ETHYL-B-NITROSTYRI
has-mm	177.2
has-stereocentres	0
has-empirical-formula	C10H11NO2
has-cas	1202-32-0

ACS 10 Sept 06

Cyber

type OrganicMolecule

has-property 48.29587

- type PhysicaProperty
- type BoilingPoint
- type Resource

has-information 49.29587

- type Resource
- type Information
- has-date 2005-2-28
- has-value 307.5

has-uncertainty

has-unit

has-source

has-method 53.29587

- type Source
- type PhysProp
- type Resource

has-property 54.29587

- type Resource
- type Laboratory

of-quality

has-condition

has-property 16.29587

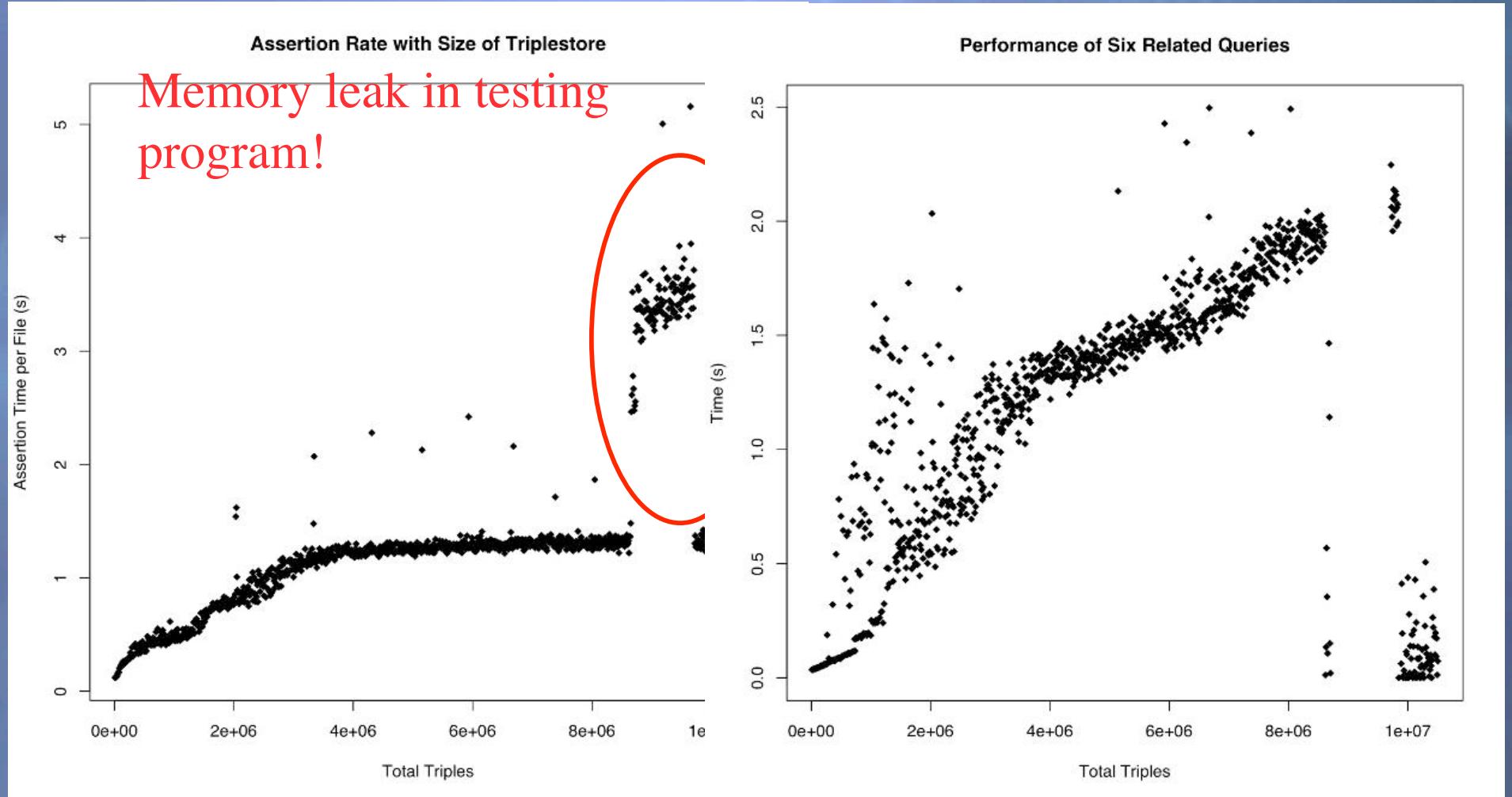
- type PhysicaProperty
- type Solubility
- type Resource

© University of Southampton 2005
ACCESS RESTRICTED
Version 0.4



Triple Stores - The Heart of the Semantic Web

Scaling - 3Store response



Scaling the triplestores

Moved from...

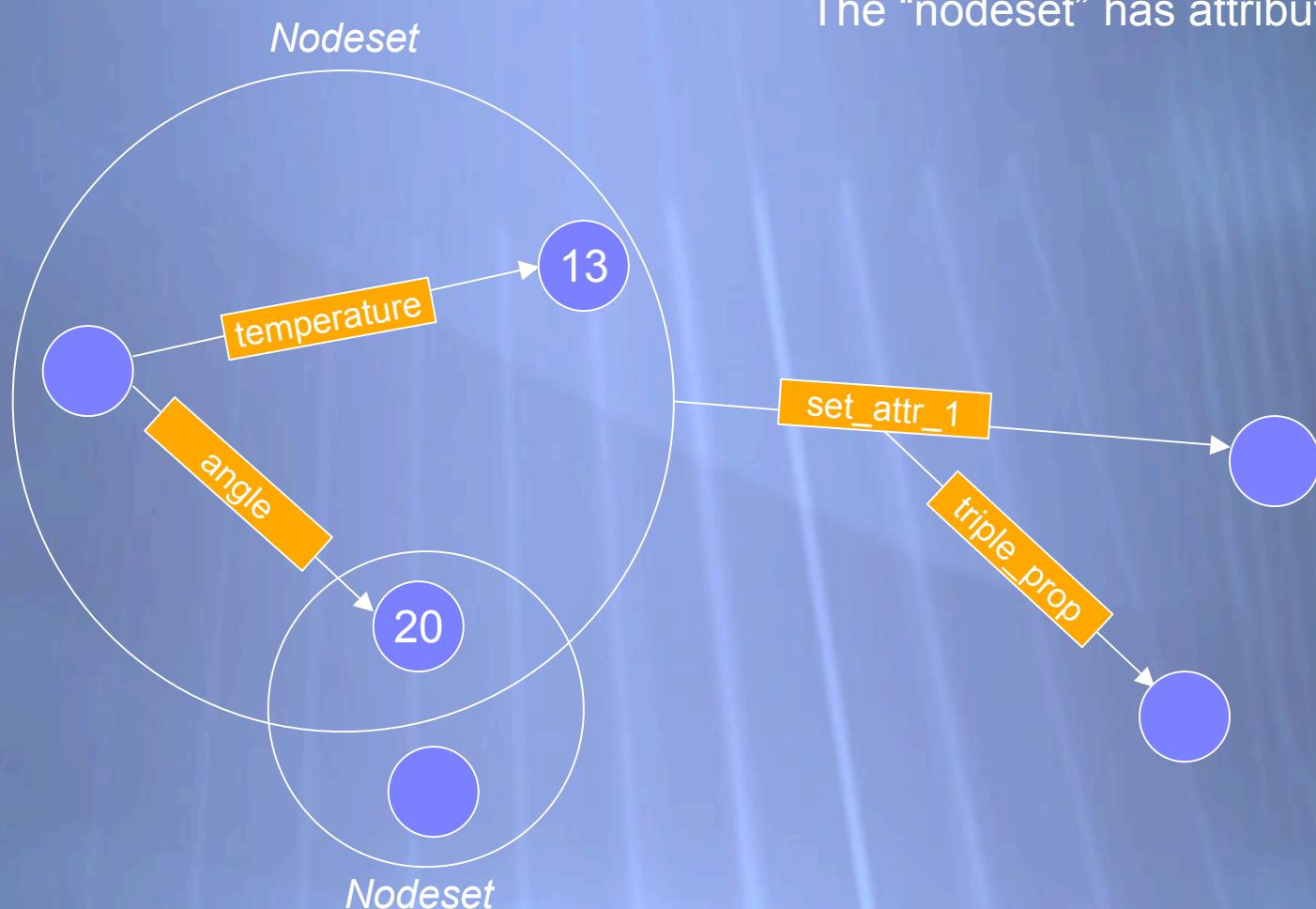
- ★ A model of harvesting data from multiple sources into one scalable store

to

- ★ A model of distributed RDF sources and caching what is needed for the task at hand into multiple stores fit-for-purpose

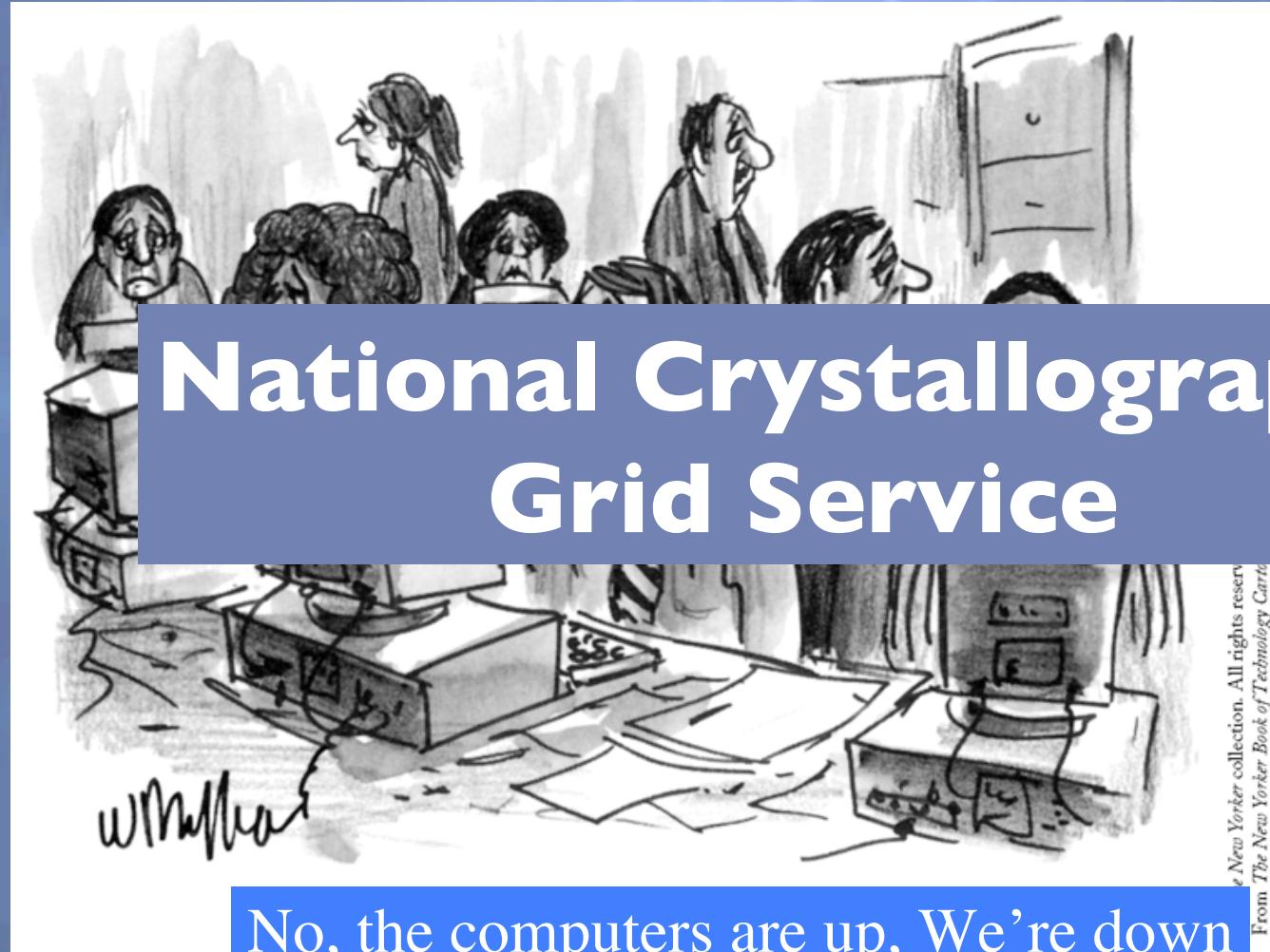
The Semantic Web!

- ★ Very Complex RDF
- ★ Use a graph-based model.
 - ★ Base on RDF ideas
 - ★ Actual data is stored as a graph
 - ★ Contrast with models like E-R, where the graph “models” the data, rather than actually being the data.
 - ★ A node in E-R might be “customer”, and represent the class of entities that are customers, rather than any specific customer.
- ★ The pseudograph model:
 - ★ Each node is a datum.
 - ★ Each edge denotes an association/attribute/property.
 - ★ Nodes can be grouped into nodesets, which are also nodes.
 - ★ A node may be in more than one nodeset.
 - ★ A node-edge-node triple can also be a node.
 - ★ Main difference from RDF is an attempt to build reification into the model.



- ★ The edge with the attribute name `set_attr_1` is an attribute of a nodeset.
- ★ The edge with the attribute name `triple_prop` is an attribute of the above edge.

Experiments on the Grid



© New Yorker collection. All rights reserved.
From The New Yorker Book of Technology Cartoons

No, the computers are up, We're down

Jeremy Frey 25

Security and
trust for
experiments
and data

Role based
authorisation
needed for
experiments
not entirely
the same as
for
computers

ACS 10 Sept 06

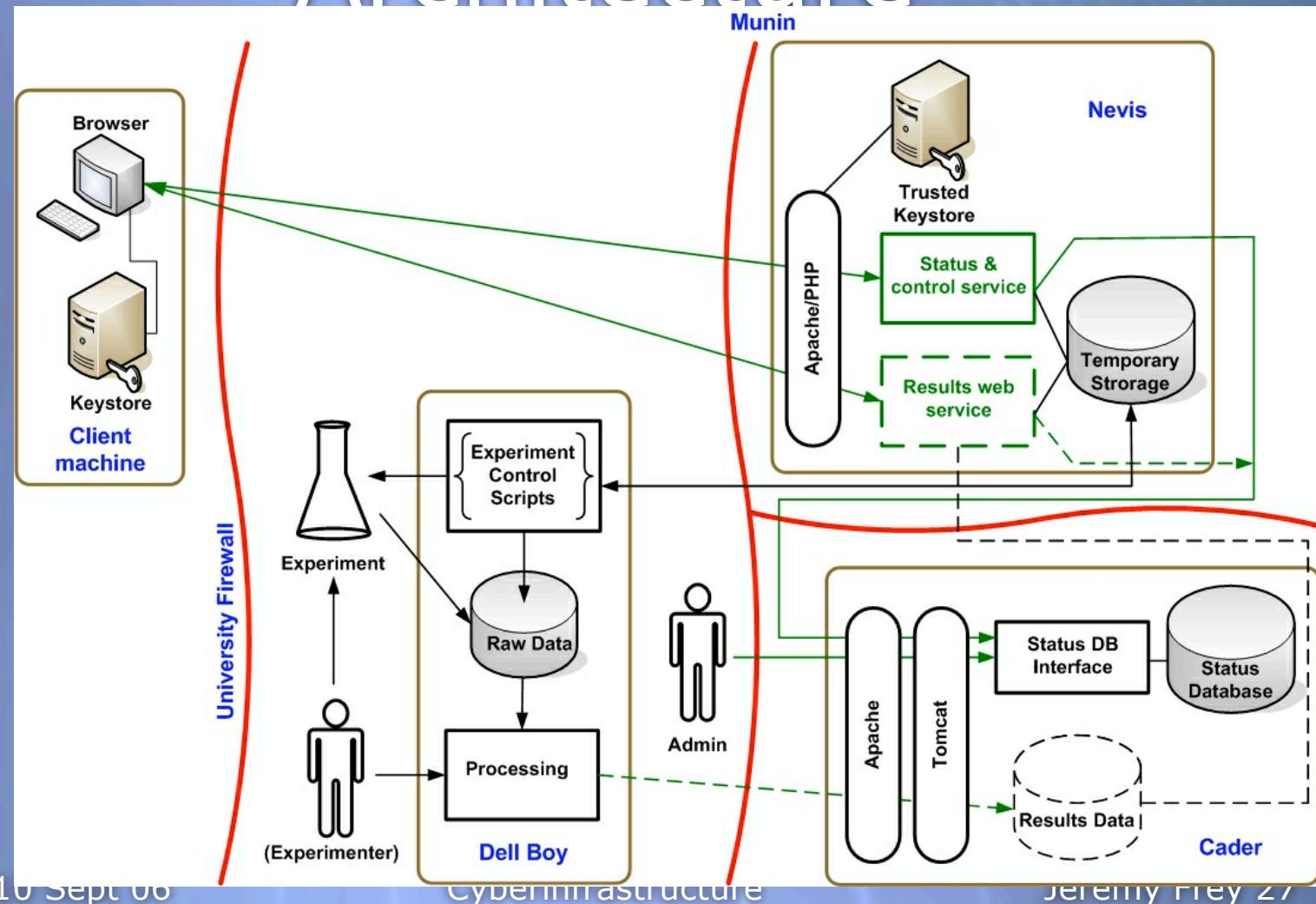


Cyberinfrastructure

Jeremy Frey 26

© The New Yorker collection. All rights reserved.
From *The New Yorker Book of Technology Cartoons*.

NCS Grid Service Architecture



Combechem status - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back

Address <https://interact.xservice.soton.ac.uk/status/index.php> Go Links >

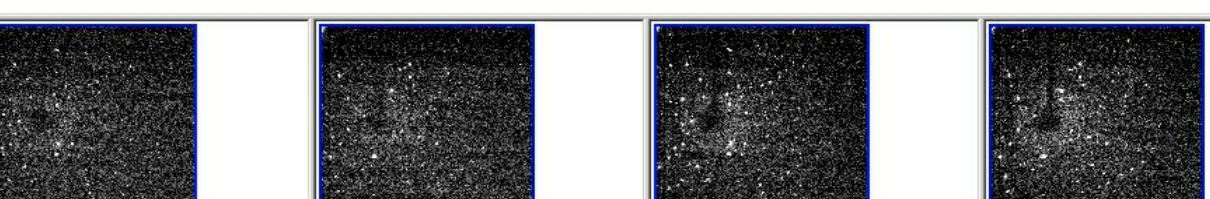
 National Crystallography Service – Sample Status

Viewing samples for M E Light (light@soton.ac.uk)

NCS ID	Customer ID	Received	Collection	Status	Details
04MEL0098	2nd test	2004-02-12	001	Succeeded	HKL file / Report
04MEL0093	mel01	2004-02-06	001	Succeeded	HKL file / Report
04SRC0104	#13-123	2004-03-08	001	Next	Due at 00:00:00 (est)
04SRC0103	#12-01	2004-03-08	001	Failed (Referred)	Diffraction too weak
			002	Failed (No Further Action)	Crystals too small
04SRC0105	HSF-HCl				

Done

X-Ray Diffractometer Images



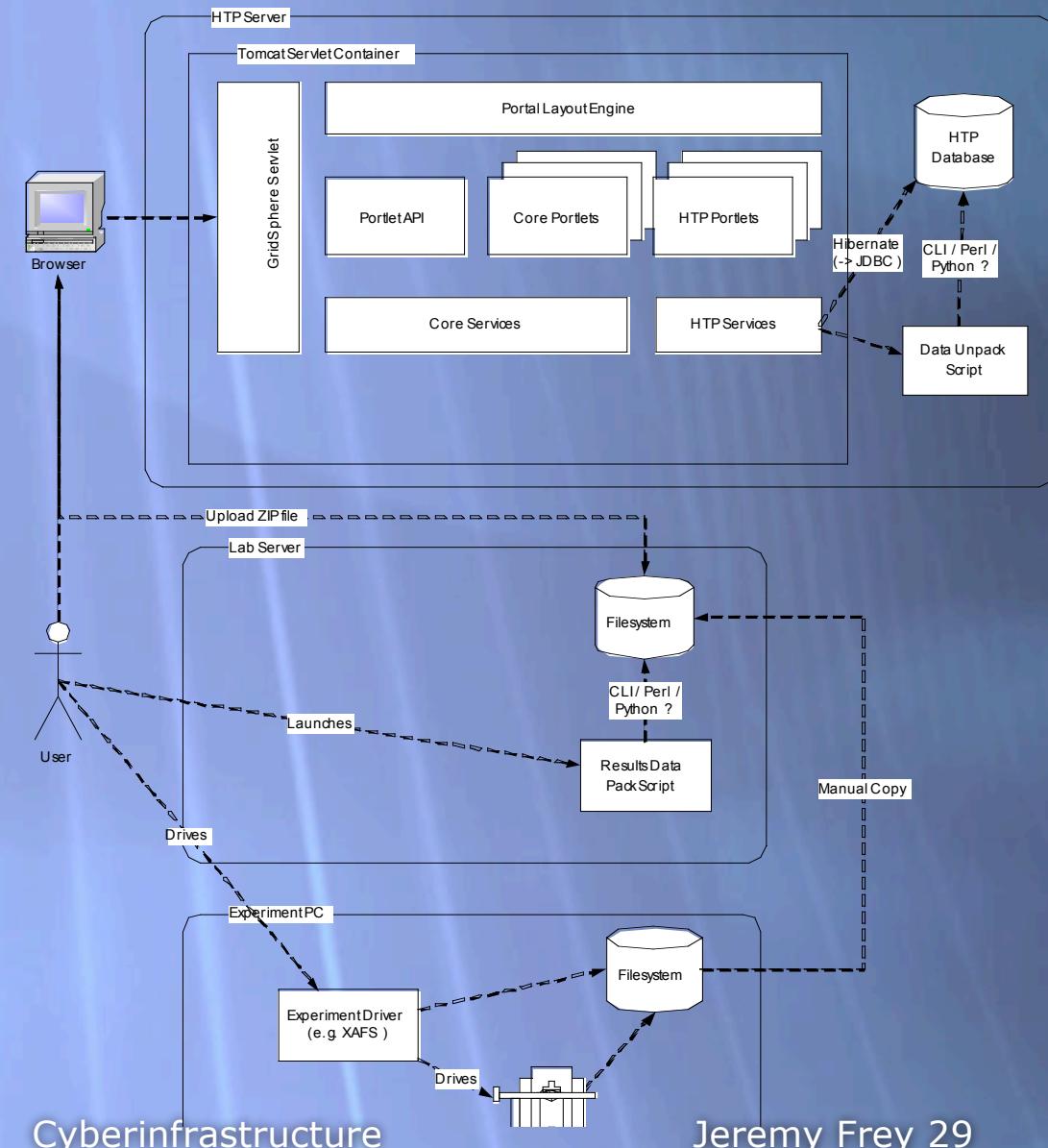
Status Log

ACS 10 September 2018 | Open Access | DOI: 10.1002/anie.201805120 | Wiley 28

HTP Sample Tracking

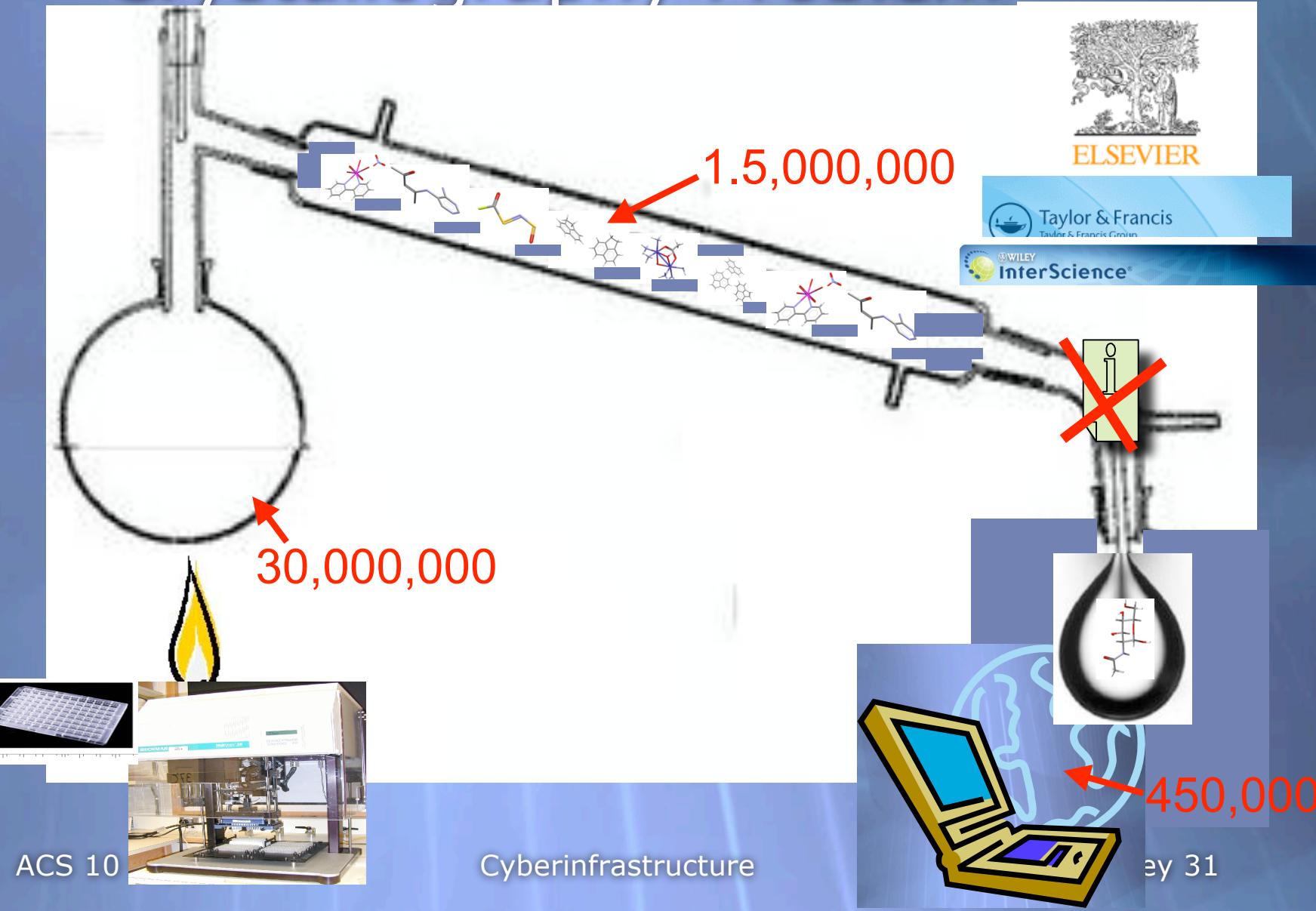
Using ideas from the NCS Grid Service we have produced a prototype for a high throughput catalyst experiment involving array samples investigated by Raman, MS, EXAFS with the samples manufactured at one site and tested at several others

HTP Architecture (First Prototype)



Dissemination

A Data-Rich Subject – the Crystallography Problem



The eCrystals Data Archive

 University of Southampton Crystal Structure Report Archive

6,7,9,10,12,13,15,16-Octahydro-benzo-1,4,7,10,13-pentaoxacyclopentadecin

Simon J Coles, Michael B Hursthouse, Jeremy G Frey and Esther Rousay.

University of Southampton

C₁₄H₂₀O₅

InChI=1/C14H20O5/c1-2-4-14-13(3-1)18-11-9-16-7-5-15-6-8-17-10-12-19-14/h1-4H,5-12H2

DOI: 10.594/ecrystals.chem.soton.ac.uk/145

Compound Class: Organic

Keywords: crown ethers

Creation Date: 07 October 2004

Deposited By: A.N. Admin

Deposited On: 20 February 2006

Available Files



Depositor Comments

Structure already known, but accurately redetermined for a local research project.

Data collection parameters

Chemical formula	C ₁₄ H ₂₀ O ₅
Crystallisation Solvent	
Crystal morphology	Plate
Crystal system	Orthorhombic
Space group symbol	Pbca
Cell length a	16.4963(18)
Cell length b	8.325(3)
Cell length c	20.061(6)
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	0.0409
R Factor (Obs)	0.0487
R Factor (All)	0.0977
Weighted R Factor (Obs)	0.1008
Weighted R Factor (All)	0.1192

Citation: Coles, S.J., Hursthouse, M.B., Frey, J.G. and Rousay, E. (2004), Southampton, UK, University of Southampton, Crystal Structure Report Archive. (doi:10.594/ecrystals.chem.soton.ac.uk/145)

Final Result

04sjc0831.cif	13k
04sjc0831.cml	6k

Validation

04sjc0831_checkcif.htm	7k
------------------------	----

Refinement

04sjc0831.res	6k
04sjc0831_xl.lst	34k

Solution

04sjc0831.prp	6k
04sjc0831_xs.lst	39k

Processing

04sjc0831.hkl	702k
04sjc0831.htm	10k
04sjc0831_0kl.jpg	57k
04sjc0831_h0l.jpg	85k
04sjc0831_hk0.jpg	88k

Data Collection

04sjc0831_crystal.jpg	17k
-----------------------	-----

Other Files

04sjc0831.doc	78k
04sjc0831.fcf.txt	155k

<http://ecrystals.chem.soton.ac.uk>

Access to the underlying data

Available Files

Data collection parameters

Chemical formula	C ₃₀ H ₂₆ FeN ₂ O ₃
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

Final Result

- 02sot064.CIF - 19k
- 02sot064.cml - 8k
- 02sot064.checkcif.html - 14k

Refinement

- 02sot064.RES - 9k

Solution

- 02sot064.PRP - 5k

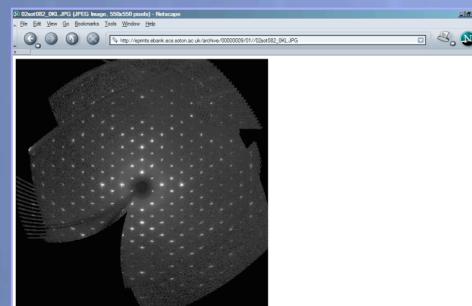
Processing

- 02SOT064.HTM - 6k
- 02sot064.HKL - 338k

Other Files

- 02sot064.DOC - 113k
- 02sot064.LST - 49k

Red arrows point from the 'Final Result' and 'Processing' sections of the main interface to the 'checkcif/PLATON report (full structural check)' and 'Data Collection Summary' windows respectively.



Act

checkcif/PLATON report (full structural check) - Netscape

checkcif/PLATON report (full structural check)

No syntax errors found.
Please wait while processing

CIF dictionary
Interpreting this report

Datablock: 1

Bond precision: C-C = 0.0952 Report set to 0.71073
Cell: a=1.2041(10) b=1.2415(10) c=10.4118(10)
alpha=97.46(10) beta=95.44(10) gamma=94.37(10)
Volume: 144.16(8)
Space group: P 1
Hall group: 1
Niggli formula: C12 H8 Cl F2 O
Min. Formula: C12 H8 Cl F2 O
Max. d: 230.65 230.65
Max. d (mm⁻¹): 1.324 1.324
Z: 2 2
Mr (mm⁻¹): 0.346 0.346
PDB: 236.0 236.0
PDB ID: 236.01 236.01
Junk: 6.15.15 6.15.15
Max: 2449 2449
Min: -2466.5 -2466 -2466.5 -2466
Theta: 0.986 0.986
Orientation method: MOSTL-OCM1
Data completeness: 100.00% Theta (max): 27.49
R(Reflections): 0.0531 (1517) wR2(Reflections): 0.1232 (2449)
S: 0.980 Rg: 198

checkcif/PLATON report (full structural check) - Netscape

checkcif/PLATON report (full structural check)

No syntax errors found.
Please wait while processing

CIF dictionary
Interpreting this report

Datablock: 1

Bond precision: C-C = 0.0952 Report set to 0.71073
Cell: a=1.2041(10) b=1.2415(10) c=10.4118(10)
alpha=97.46(10) beta=95.44(10) gamma=94.37(10)
Volume: 144.16(8)
Space group: P 1
Hall group: 1
Niggli formula: C12 H8 Cl F2 O
Min. Formula: C12 H8 Cl F2 O
Max. d: 230.65 230.65
Max. d (mm⁻¹): 1.324 1.324
Z: 2 2
Mr (mm⁻¹): 0.346 0.346
PDB: 236.0 236.0
PDB ID: 236.01 236.01
Junk: 6.15.15 6.15.15
Max: 2449 2449
Min: -2466.5 -2466 -2466.5 -2466
Theta: 0.986 0.986
Orientation method: MOSTL-OCM1
Data completeness: 100.00% Theta (max): 27.49
R(Reflections): 0.0531 (1517) wR2(Reflections): 0.1232 (2449)
S: 0.980 Rg: 198

checkcif/PLATON report (full structural check) - Netscape

checkcif/PLATON report (full structural check)

No syntax errors found.
Please wait while processing

CIF dictionary
Interpreting this report

Datablock: 1

Bond precision: C-C = 0.0952 Report set to 0.71073
Cell: a=1.2041(10) b=1.2415(10) c=10.4118(10)
alpha=97.46(10) beta=95.44(10) gamma=94.37(10)
Volume: 144.16(8)
Space group: P 1
Hall group: 1
Niggli formula: C12 H8 Cl F2 O
Min. Formula: C12 H8 Cl F2 O
Max. d: 230.65 230.65
Max. d (mm⁻¹): 1.324 1.324
Z: 2 2
Mr (mm⁻¹): 0.346 0.346
PDB: 236.0 236.0
PDB ID: 236.01 236.01
Junk: 6.15.15 6.15.15
Max: 2449 2449
Min: -2466.5 -2466 -2466.5 -2466
Theta: 0.986 0.986
Orientation method: MOSTL-OCM1
Data completeness: 100.00% Theta (max): 27.49
R(Reflections): 0.0531 (1517) wR2(Reflections): 0.1232 (2449)
S: 0.980 Rg: 198

checkcif/PLATON report (full structural check) - Netscape

checkcif/PLATON report (full structural check)

No syntax errors found.
Please wait while processing

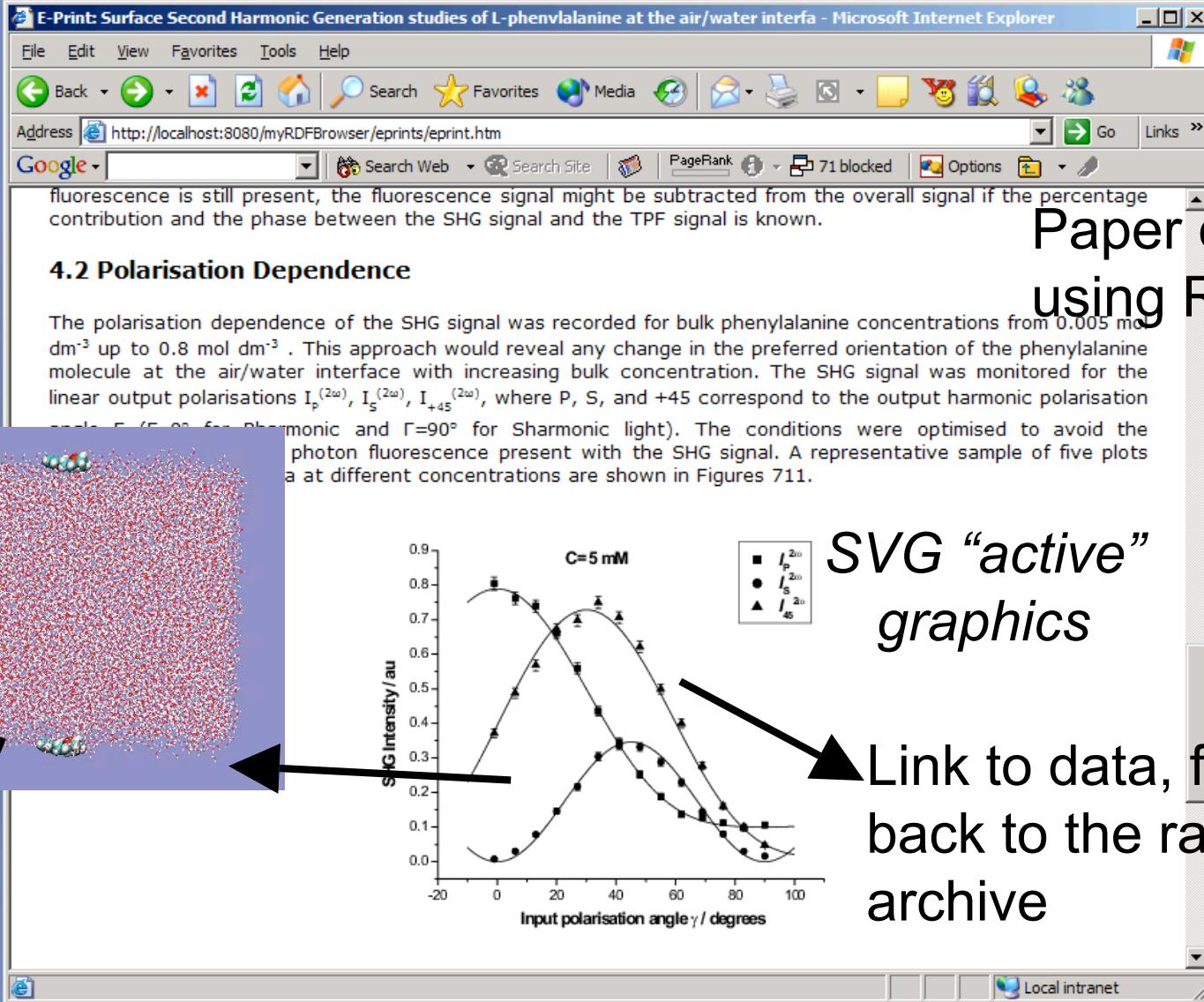
CIF dictionary
Interpreting this report

Datablock: 1

Bond precision: C-C = 0.0952 Report set to 0.71073
Cell: a=1.2041(10) b=1.2415(10) c=10.4118(10)
alpha=97.46(10) beta=95.44(10) gamma=94.37(10)
Volume: 144.16(8)
Space group: P 1
Hall group: 1
Niggli formula: C12 H8 Cl F2 O
Min. Formula: C12 H8 Cl F2 O
Max. d: 230.65 230.65
Max. d (mm⁻¹): 1.324 1.324
Z: 2 2
Mr (mm⁻¹): 0.346 0.346
PDB: 236.0 236.0
PDB ID: 236.01 236.01
Junk: 6.15.15 6.15.15
Max: 2449 2449
Min: -2466.5 -2466 -2466.5 -2466
Theta: 0.986 0.986
Orientation method: MOSTL-OCM1
Data completeness: 100.00% Theta (max): 27.49
R(Reflections): 0.0531 (1517) wR2(Reflections): 0.1232 (2449)
S: 0.980 Rg: 198

Cyber

Jeremy Frey 33



Paper organized using RDF

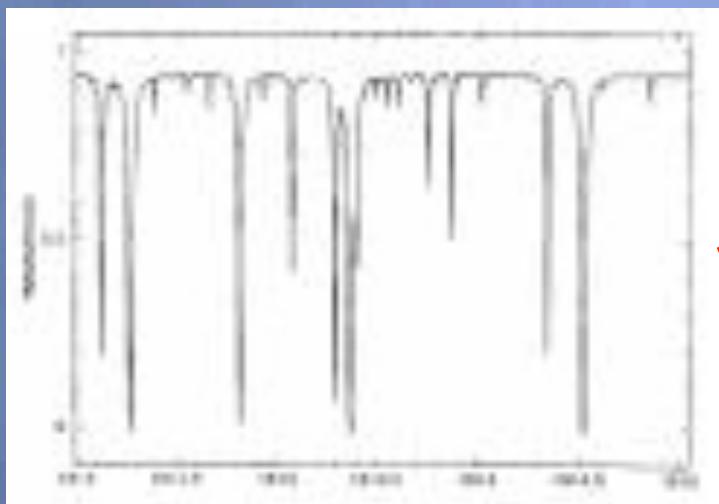
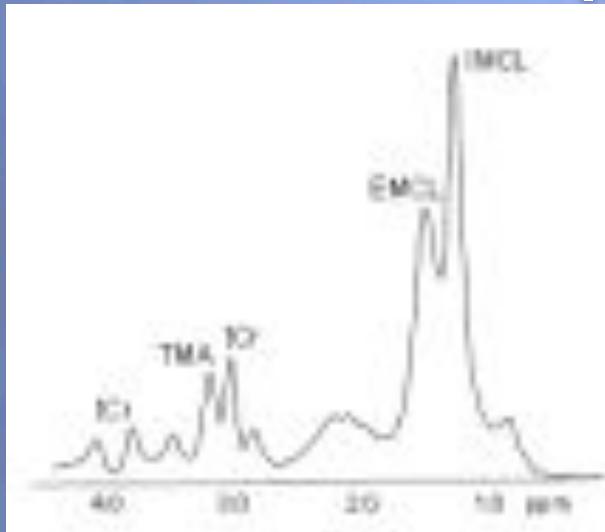
Link to simulation, full simulation data archived in BioSimGrid

SVG “active” graphics

Link to data, follow links back to the raw data archive

R4L

Need for a data archive in the laboratory



b515548g.pdf (application/pdf Object) - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

NC5 Interactive Serv... Repositories 4 Labs ... Calendar WWWgetdata

NC5 Dalton Transactions Articles b515548g.pdf (application/pdf Object)

Save a Copy Select Search Web YAHOO! TOOLBAR

Preparation of $\text{CPh}_3[\text{C}_6\text{F}_5]_3\text{BCNPBB}$ (2)

by passing through columns of P_2O_5 with moisture indicator and 4 Å molecular sieves. Gel permeation chromatography (GPC) measurements were performed on a Polymer Laboratories PL-GPC-220 instrument equipped with a $\text{W}5\text{L}5\text{Å}$ Mixed-C column, a refractive index detector, and a PD2040 light scattering detector. The GPC column was calibrated using eight monodisperse polystyrene standards in the range 580–483000 Da.

Preparation of $\text{CPh}_3[\text{C}_6\text{F}_5]_3\text{BCNPBB}$ (1)

Potassium cyanide (33.6 mg, 0.5 mmol) was ground to a powder using a pestle and mortar in a dry box. PBB (0.478 g, 0.5 mmol) and 50 mL diethyl ether were then added, and the mixture was heated to reflux for 12 h. The solvent was removed *in vacuo* to leave an off-white foam which was washed with warm hexane (50 mL) to give K[NCPBB] as a white powder (0.495 g, 485 mmol). This solid was stirred with triphenylchloromethane (0.135 g, 0.485 mmol) in dichloromethane (15 mL) for 20 h. The solution was filtered to remove KCl, concentrated to *ca.* 5 mL, and cooled to $-26\text{ }^\circ\text{C}$ to give an orange crystalline solid, yield: 0.324 g (0.315 mmol, 63% with respect to KCN). IR (nujol): 2189 cm^{-1} ($\nu_{\text{C}\equiv\text{C}}$). ^1H NMR (CD_2Cl_2 , $20\text{ }^\circ\text{C}$, 300.13 MHz): δ 8.28 (t, 3 H, $J = 7.5\text{ Hz}$, *p*-Ph), 7.70 (t, 6 H, $J = 7.7\text{ Hz}$, *m*-Ph), 7.70 (d, 6 H, $J = 7.3\text{ Hz}$, *o*-Ph). ^{13}C NMR (CD_2Cl_2 , $20\text{ }^\circ\text{C}$, 75.48 MHz): δ 211.4 (CPh_3), 144.1 (*p*-C), 143.0 (*m*-C), 140.1 (*ipso*-C), 130.9 (*o*-C). ^{11}B NMR (CD_2Cl_2 , $20\text{ }^\circ\text{C}$, 96.3 MHz): δ -4.35 (s, br, 1 B, N- $\text{B}(\text{C}_6\text{F}_5)_3$), -18.27 (s, 1 B, C- $\text{B}(\text{C}_6\text{F}_5)_3$). ^{19}F NMR (CD_2Cl_2 , $20\text{ }^\circ\text{C}$, 282.4 MHz): δ -18.72 (br, s, 1 F), -120.22 (br, s, 1 F), -121.99 (br, s, 1 F), -122.50 (s, 1 F), -132.20 (s, 1 F), -133.94 (br, 6 F, *o*-F on $\text{B}(\text{C}_6\text{F}_5)_3$), overlapping signals (-134.15, -134.39, -134.95, -135.27, -135.64), 136.89 (br, 1 F), -137.81 (br, 3 F), -138.79 (d, 1 F), -144.73 (t, 1 F), -149.78 (t, 1 F), -151.11 (t, 1 F), -154.65 (t, 1 F), -154.93 (t, 1 F), -155.32 (t, 1 F), -156.86 (t, 1 F), -157.24 (t, 1 F), -157.55 (t, 1 F), -158.29 (t, 1 F), -158.90 (t, 1 F), -159.57 (t, 3 F, $J = 20\text{ Hz}$, *p*-F on $\text{B}(\text{C}_6\text{F}_5)_3$), -159.98 (t, 1 F), -161.44 (br, 2 F), -164.0 to -164.4 (overlapping signals, 3 F), -165.33 (br, 2 F), -166.12 (t, 3 F, $J = 20\text{ Hz}$, *m*-F on $\text{B}(\text{C}_6\text{F}_5)_3$). Anal. Calcd for $\text{C}_{24}\text{H}_{15}\text{B}_2\text{F}_{15}\text{N}$:

Done 9 of 12

Start Current WinGX v1.70.01 : Cry... Microsoft PowerPoint - [J... b515548g.pdf (appli... 15:51

Not just the published spectra!
Cyberinfrastructure Jeremy Frey 35

The R4L Repository Deposit

Chemical Name *
The name of the chemical compound which is the subject of the experiment.

InChI
InChI (International Chemical Identifier) code describing the chemical compound which is the subject of the experiment.

Experimental Technique *
The technique used to produce data about the chemical compound.

Single Crystal Diffraction
Powder X-Ray Diffraction
Infra-Red Spectroscopy
Ultra-Violet Spectroscopy
Mass Spectroscopy
Raman Spectroscopy
Optical Microscopy
Differential Scanning Calorimetry
Thermogravimetric Analysis
Nuclear Magnetic Resonance

Instrument *
Manufacturer and model of instrument used.

Date *
The date that the experiment was carried out.
"Day" or both "Month" and "Day" may be omitted if appropriate.

Year: Month: Unspecified Day:

Time *
The time that the experiment was carried out.

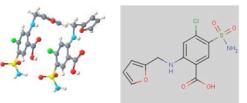
Experimenters *
The research worker(s) who conducted the experiment.

Family Name	Given Name / Initials	Email (if known)
1. <input type="text"/>	<input type="text"/>	<input type="text"/>
2. <input type="text"/>	<input type="text"/>	<input type="text"/>
3. <input type="text"/>	<input type="text"/>	<input type="text"/>
4. <input type="text"/>	<input type="text"/>	<input type="text"/>
More Spaces		



R4L Repository for the Laboratory

Furosemide Form1
Furosemide Form1 (Single Crystal Diffraction): 10 December 2003 14:05. Hughes, David S and Karami, Sohrab (for: Hursthouse, Michael B and Coles, Simon J).



Experiment data files:

- [furosemide1.sif \(955 kB\)](#)
- [furosemide1.cif \(8 kB\)](#)
- [furosemide1.mif \(8 kB\)](#)
- [furosemide1.mos \(9 kB\)](#)
- [furosemide1.pdf \(25 kB\)](#)
- [furosemide1.tif \(2 kB\)](#)
- [furosemide1.xml \(2 kB\)](#)
- [furosemide1.mol \(3 kB\)](#)

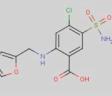
Instrument: Bruker-Netus KappaCCD Area Detector
ID: Code: 6
Deposited: 06 October 2005
By: [Simon Coles](#)
Deposited- 06 October 2005
On:

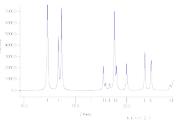
[Repository Staff Only: edit this item](#)

Deposit

R4L Repository for the Laboratory

Furosemide Form1
Furosemide Form1 (Powder X-Ray Diffraction): 14 September 2004 10:19. Hughes, David S and Light, Mark E (for: Hursthouse, Michael B and Coles, Simon J).





Experiment data files:

- [furosemide1.dat \(17 kB\)](#)
- [furosemide1.raw \(7 kB\)](#)
- [furosemide1.mif \(3 kB\)](#)
- [index.html \(1 kB\)](#)
- [furosemide1.mol \(3 kB\)](#)

InChI: [InChI=1/C12H11ClN2O5S/c2*13-9-5-10\(15-6-7-2-1-3-20-7\)8\(12\(16\)17\)4-11\(9\)21\(14,18\)19/h2*1-5,15H,6H2,\(H,16,17\)\(H2,14,18,19\)](#)
Instrument: Bruker D8000
ID: Code: 7
Deposited: [Simon Coles](#)
By: [Simon Coles](#)
Deposited: 06 October 2005
On:

[Repository Staff Only: edit this item](#)

[Contact Information](#)



R4L Repository for the Laboratory

Home About Browse Search Register User Area Help

InChI:
[InChI=1/C12H11ClN2O5S/c2*13-9-5-10\(15-6-7-2-1-3-20-7\)8\(12\(16\)17\)4-11\(9\)21\(14,18\)19/h2*1-5,15H,6H2,\(H,16,17\)\(H2,14,18,19\)](#)

Number of items: 4.

[Furosemide Form1](#) (Optical Microscopy): 30 September 2004 15:57. Hughes, David S (for: Hursthouse, Michael B and Coles, Simon J).

[Furosemide Form1](#) (Powder X-Ray Diffraction): 14 September 2004 10:19. Hughes, David S and Light, Mark E (for: Hursthouse, Michael B and Coles, Simon J).

[Furosemide Form1](#) (Infra-Red Spectroscopy): 29 July 2004 12:09. Hughes, David S (for: Hursthouse, Michael B and Coles, Simon J).

[Furosemide Form1](#) (Single Crystal Diffraction): 10 December 2003 14:05. Hughes, David S and Karami, Sohrab (for: Hursthouse, Michael B and Coles, Simon J).

This list was generated on **Sun Nov 6 19:30:03 GMT 2005**.

[Contact Information](#)

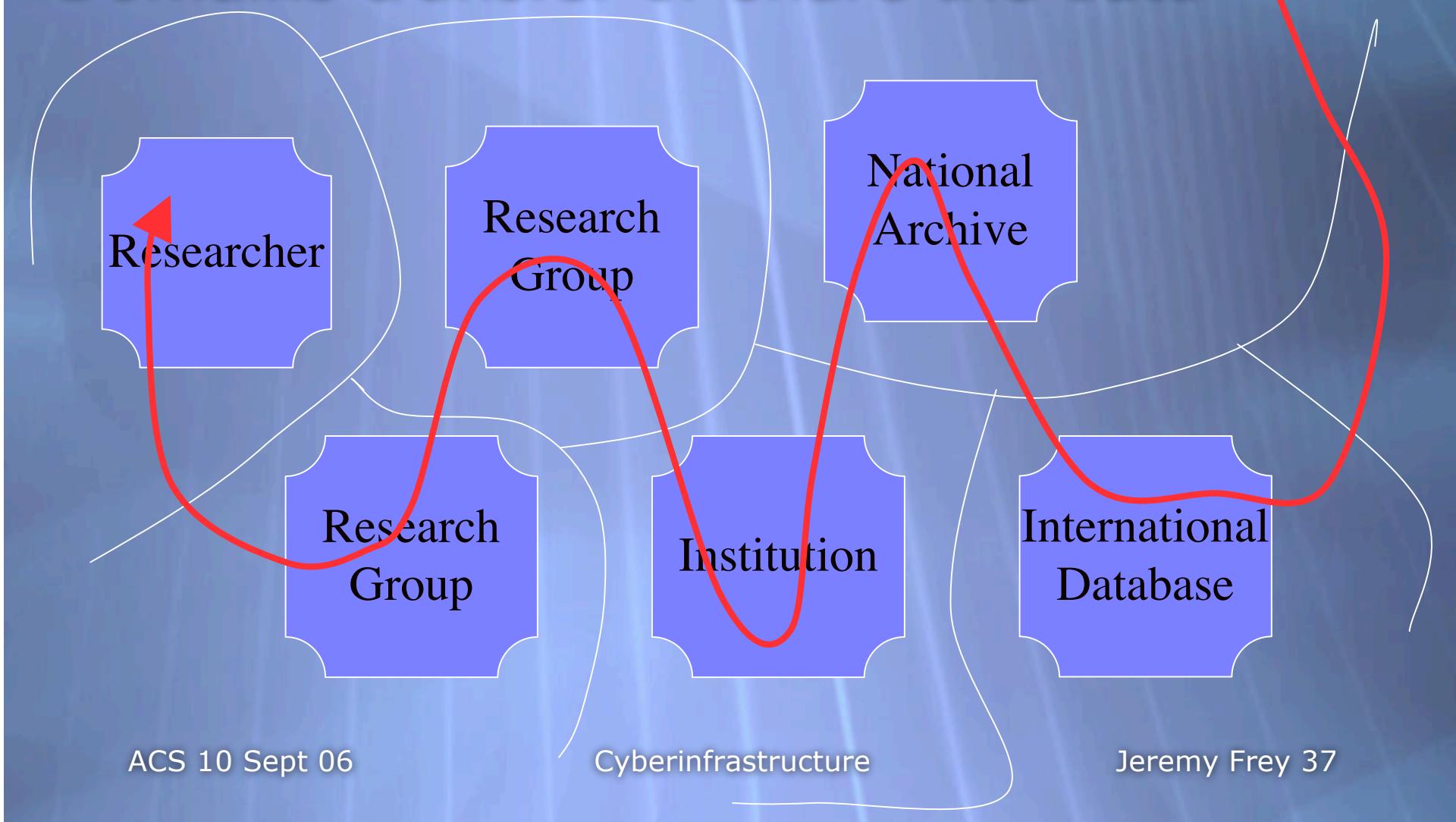
7/25 18 Sept 05

Cyberinfrastructure

Search

Jeremy Frey 36

Several groups making and analysing the library Administrative Domains transfer or share the data

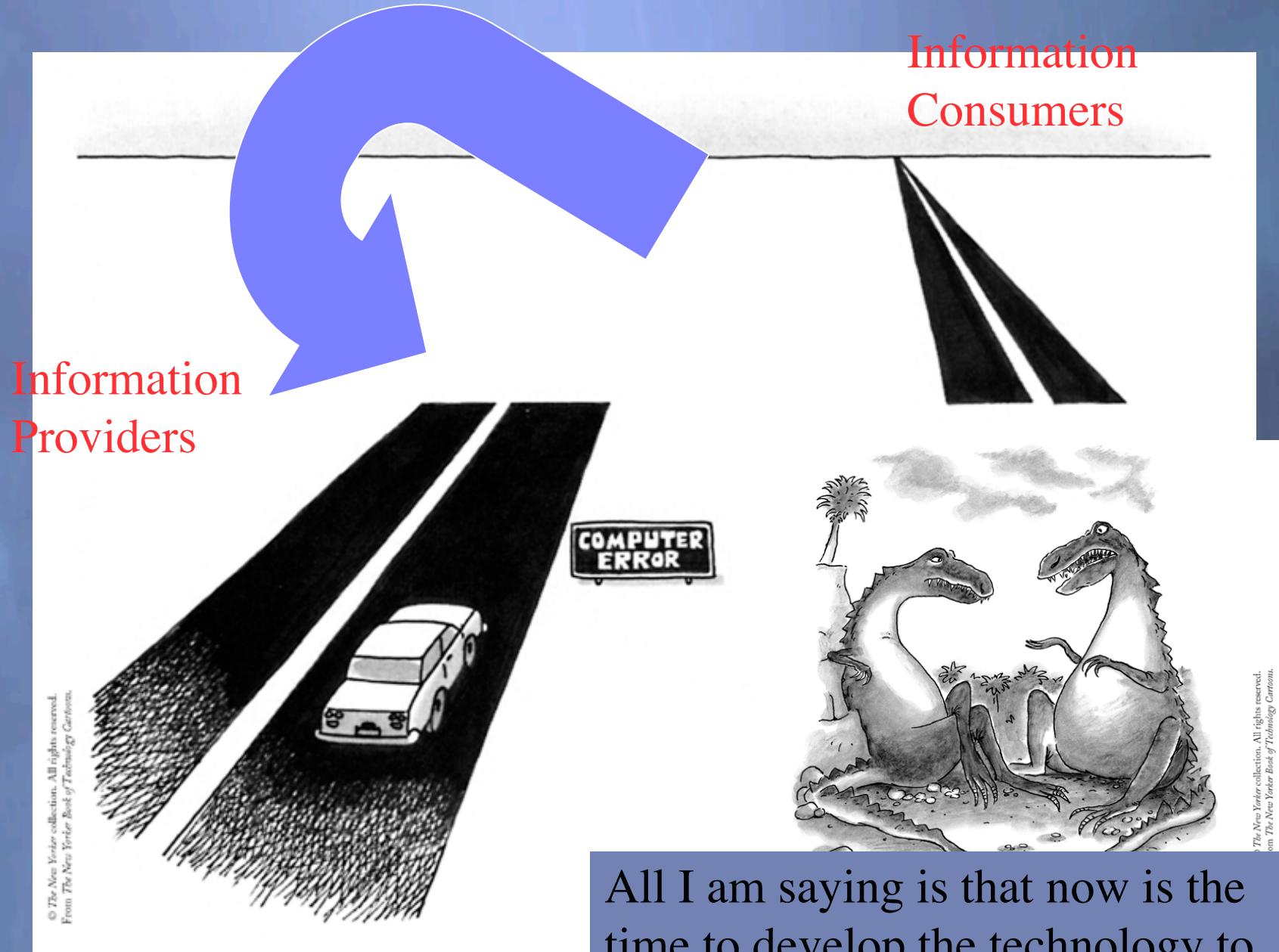


Summary:

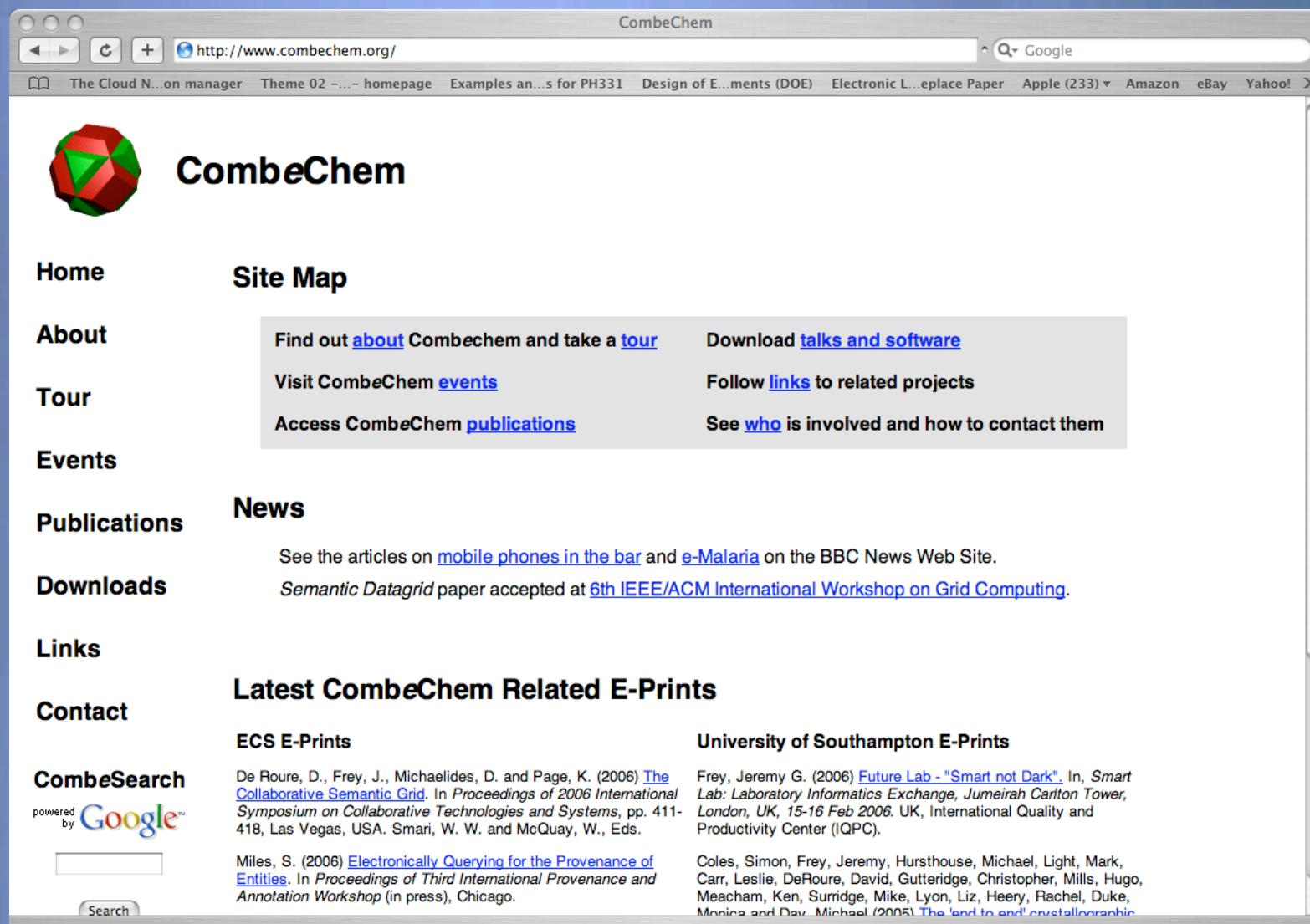
- ★ Making sure other people can find, understand and re-use your data easily and with confidence (even when there is a huge amount of it!)
- ★ Make use of Plans to inform the digital context - metadata in advance
- ★ Have concern for the “End-to-End life cycle” of chemistry information from the start.
- ★ Understanding Usability and Human Computer Interaction is vital for adoption

Summary:

- ★ Is the semantic web/grid enough
- ★ It is sufficient to evolve and improve the practice of chemistry
- ★ It is not enough - 'Pictures are worth a thousand words' more like 10,000 words of metadata
- ★ The Semiotic Web will be needed to improve the communication of models



All I am saying is that now is the time to develop the technology to deflect an asteroid



The screenshot shows a web browser window for the CombeChem website. The address bar shows the URL <http://www.combechem.org/>. The browser's toolbar includes standard icons for back, forward, and search, with a Google search bar. The page content is as follows:

CombeChem

CombeChem logo (a red, green, and blue geometric shape)

Site Map

- Home**
- About**
- Tour**
- Events**
- Publications**
- Downloads**
- Links**
- Contact**

CombeSearch powered by **Google**

Site Map

Find out <u>about</u> Combechem and take a <u>tour</u>	Download <u>talks and software</u>
Visit CombeChem <u>events</u>	Follow <u>links</u> to related projects
Access CombeChem <u>publications</u>	See <u>who</u> is involved and how to contact them

News

See the articles on [mobile phones in the bar](#) and [e-Malaria](#) on the BBC News Web Site.
Semantic Datagrid paper accepted at [6th IEEE/ACM International Workshop on Grid Computing](#).

Latest CombeChem Related E-Prints

ECS E-Prints	University of Southampton E-Prints
De Roure, D., Frey, J., Michaelides, D. and Page, K. (2006) The Collaborative Semantic Grid . In <i>Proceedings of 2006 International Symposium on Collaborative Technologies and Systems</i> , pp. 411-418, Las Vegas, USA. Smart, W. W. and McQuay, W., Eds.	Frey, Jeremy G. (2006) Future Lab - "Smart not Dark" . In, <i>Smart Lab: Laboratory Informatics Exchange, Jumeirah Carlton Tower, London, UK, 15-16 Feb 2006</i> . UK, International Quality and Productivity Center (IQPC).
Miles, S. (2006) Electronically Querying for the Provenance of Entities . In <i>Proceedings of Third International Provenance and Annotation Workshop</i> (in press), Chicago.	Coles, Simon, Frey, Jeremy, Hursthouse, Michael, Light, Mark, Carr, Leslie, DeRoure, David, Gutteridge, Christopher, Mills, Hugo, Meacham, Ken, Surridge, Mike, Lyon, Liz, Heery, Rachel, Duke, Monica and Day, Michael (2005) The 'end to end' crystallographic