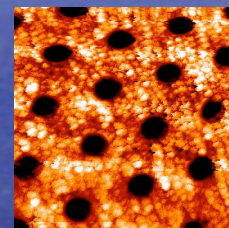
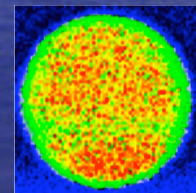


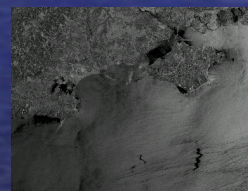
**X-ray single
Mol**



STM

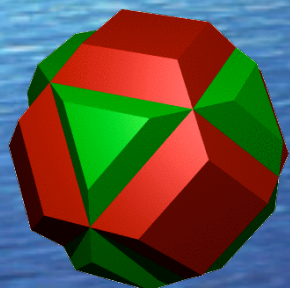


Raman



**Ocean
Monolayer**

From e-Science to Publication@Source



Comb-e-Chem

Jeremy Frey

School of Chemistry

University of Southampton, UK

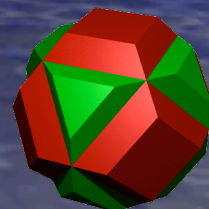
e-Science

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

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.'

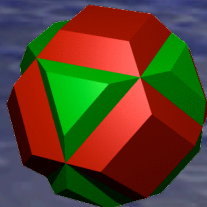
Tony Blair, 2002



The Collaboratory Concept

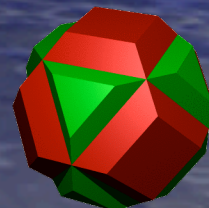
- In 1989, William Wulf, then with the U.S. National Science Foundation, defined a collaboratory as

"a center without walls, in which the nation's researchers can perform their research without regard to geographical location, interacting with colleagues, accessing instrumentation, sharing data and computational resources, and accessing information in digital libraries."

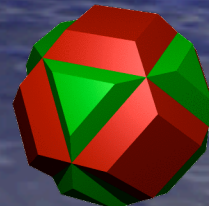
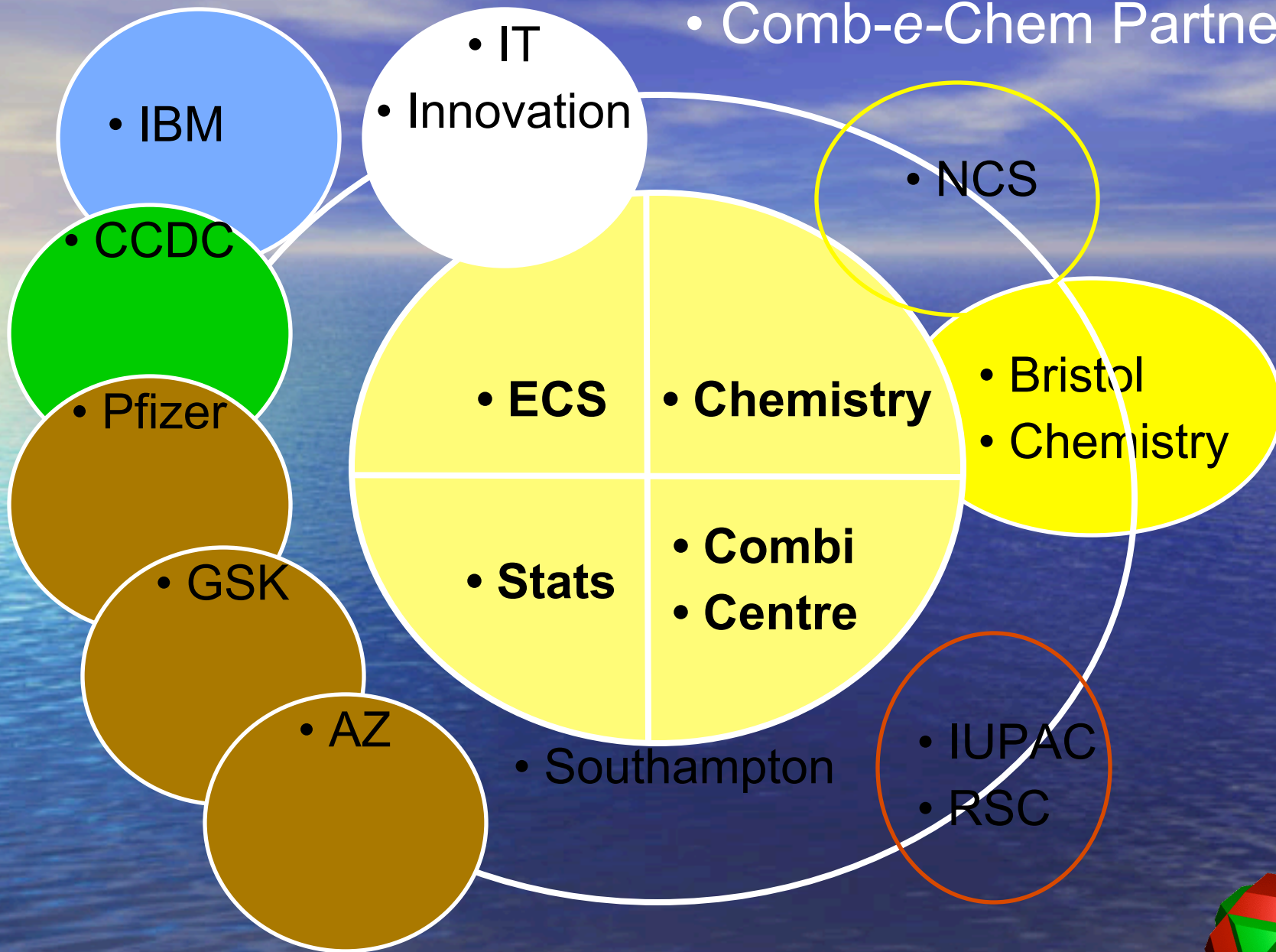


The Comb-e-Chem Project

- The exponential world of Combinatorial Synthesis and High throughput analysis meets the exponentially growing power of computing

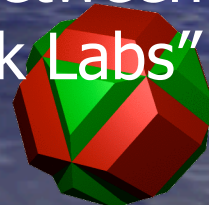
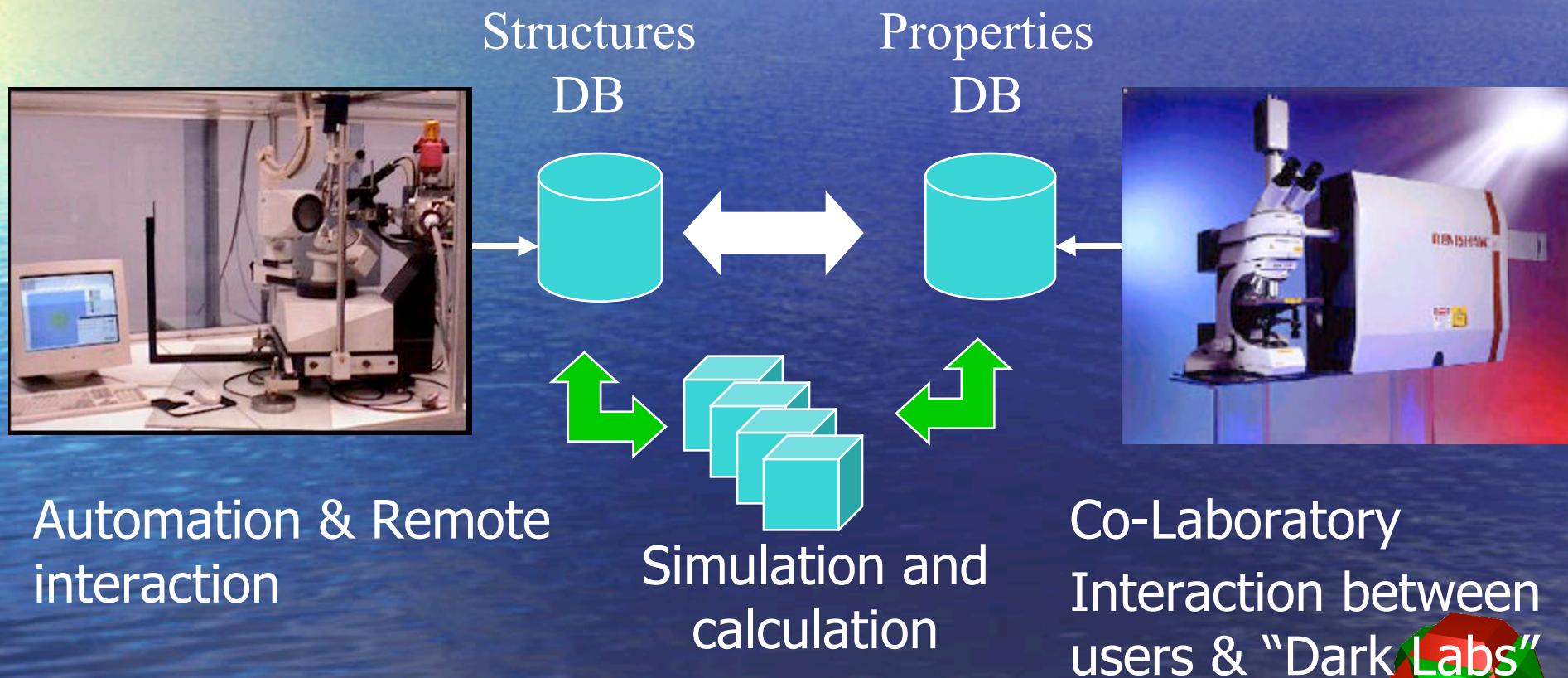


• Comb-e-Chem Partners

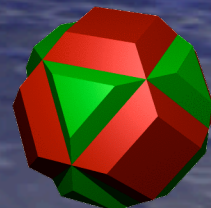
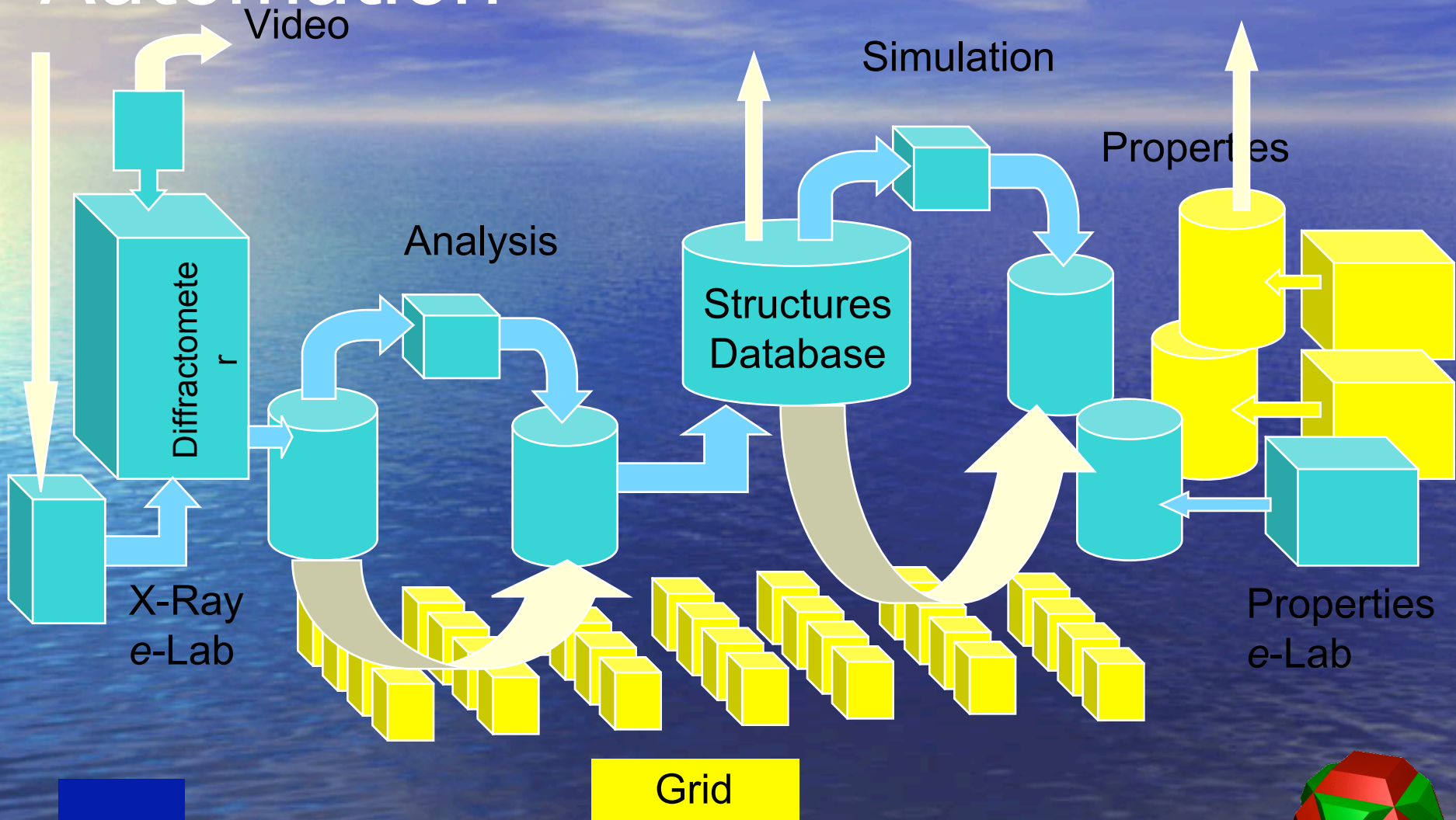


The Comb-e- Chem Vision

Structure + Properties \rightleftharpoons Knowledge + Prediction

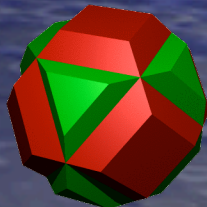
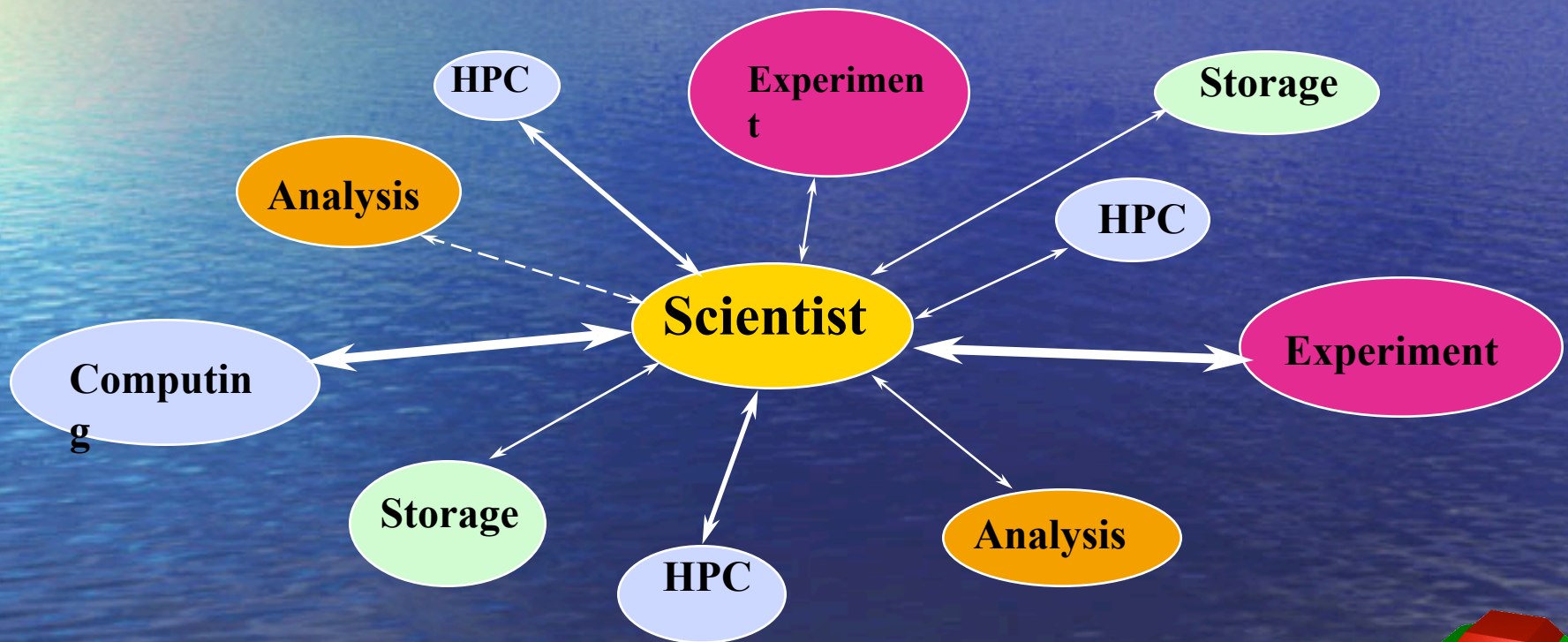


Comb-e-Chem Project - Automation



Scientist at the Centre of an Information Web

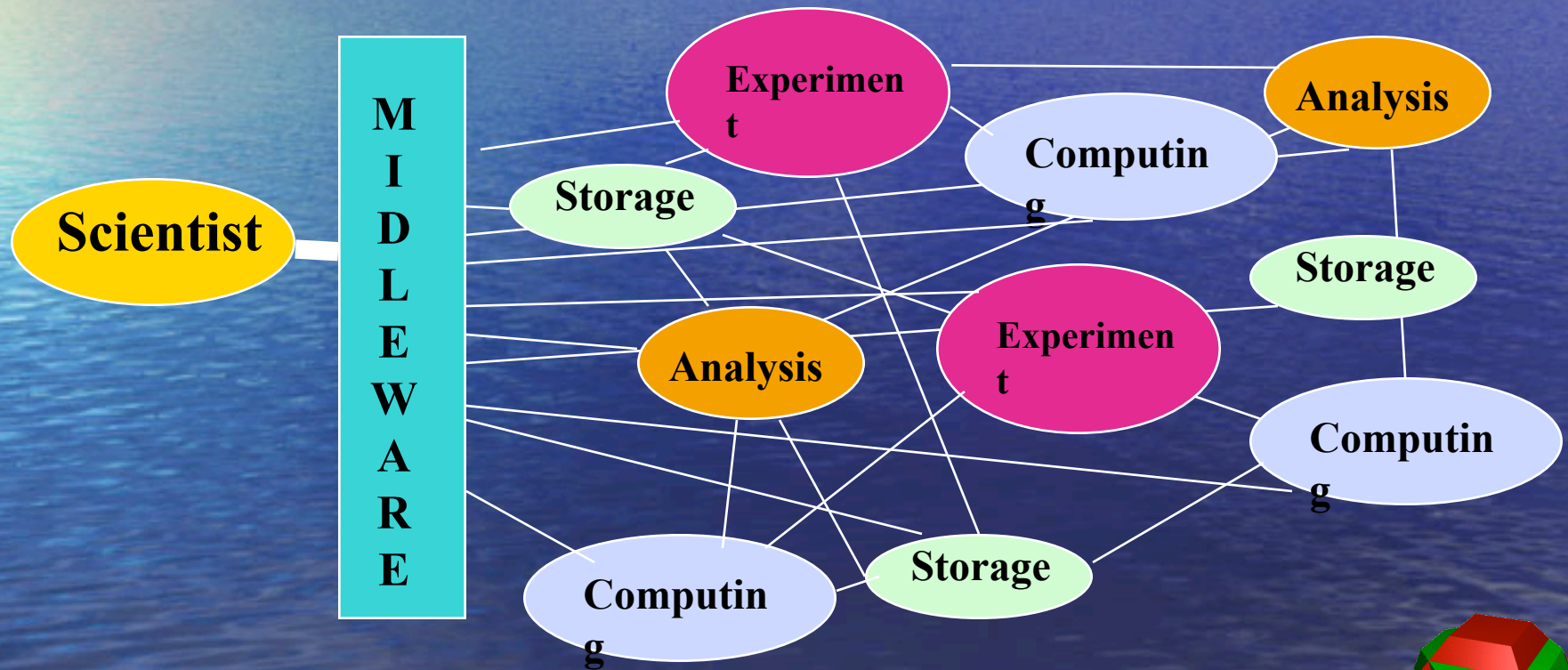
By access variable and difficult



The Future

The Grid Model - *Information Utilities*

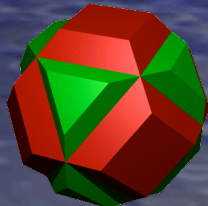
Uniform access



Remember that you contribute to other people's information web

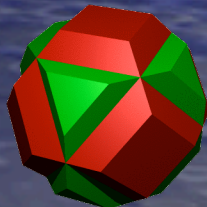
19 Feb 2004 OAI Meeting

Jeremy G. Frey & Mike Hursthouse

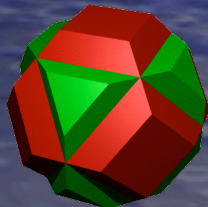
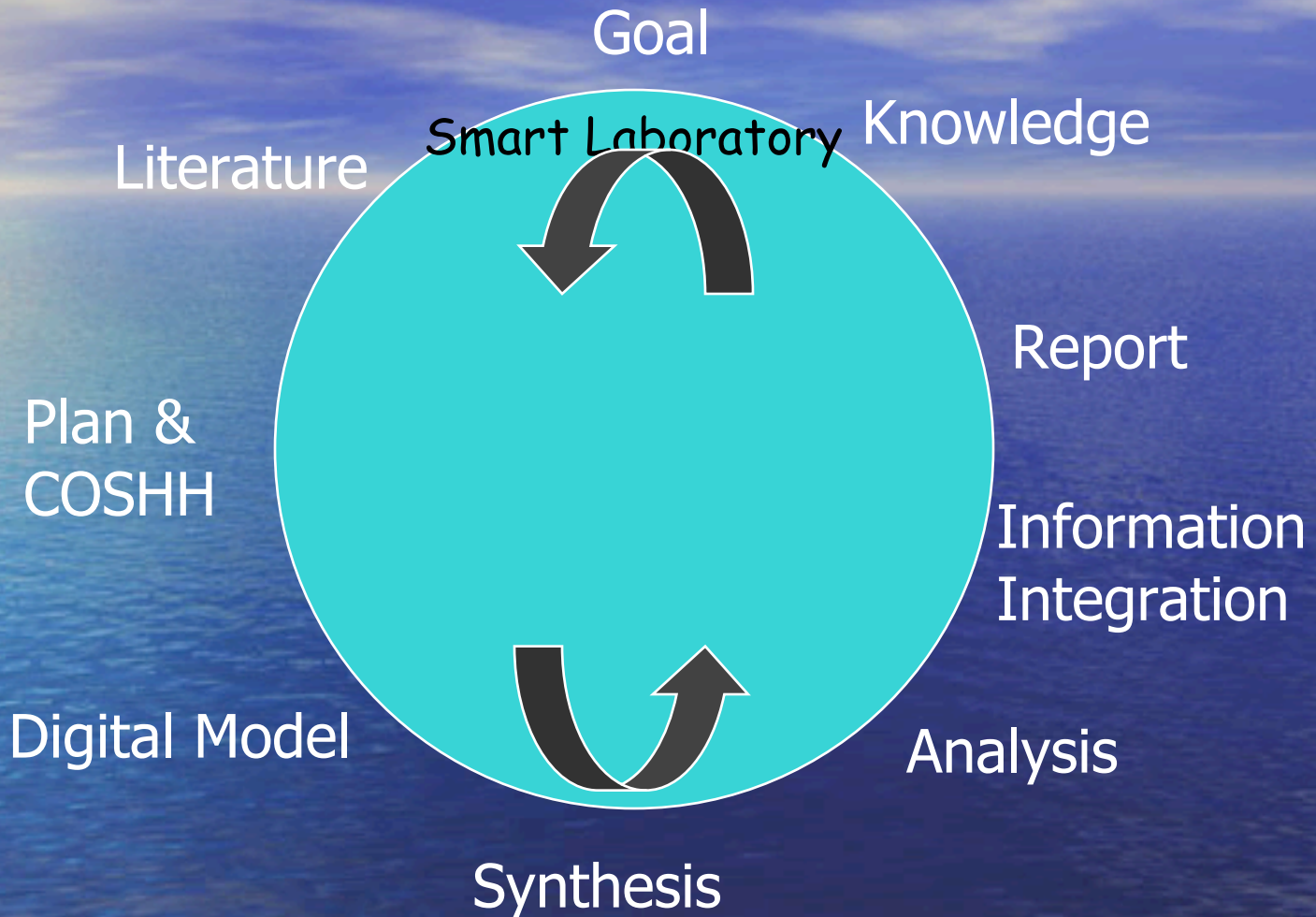


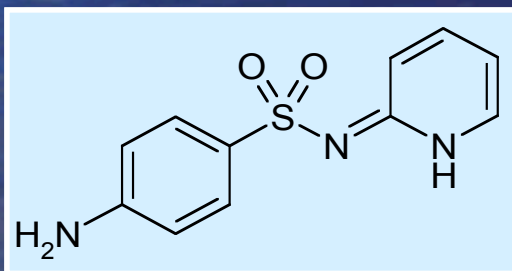
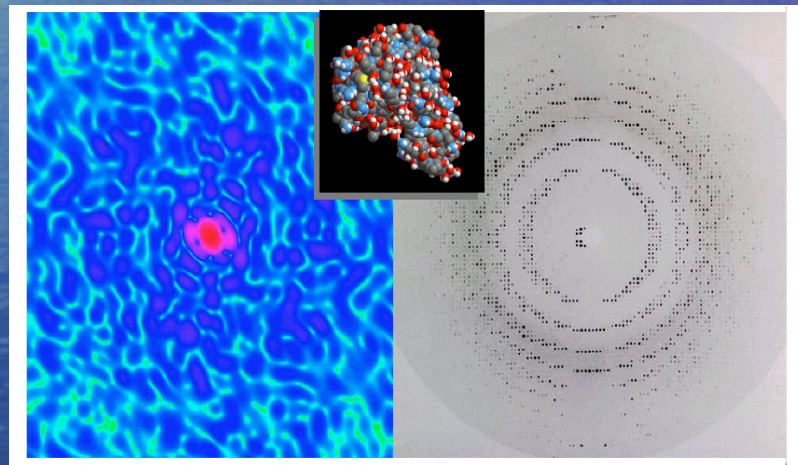
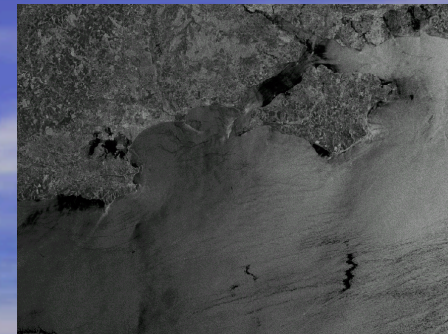
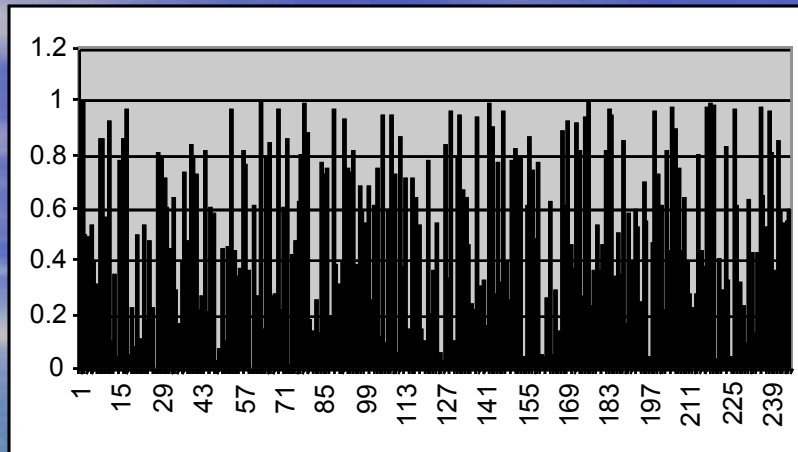
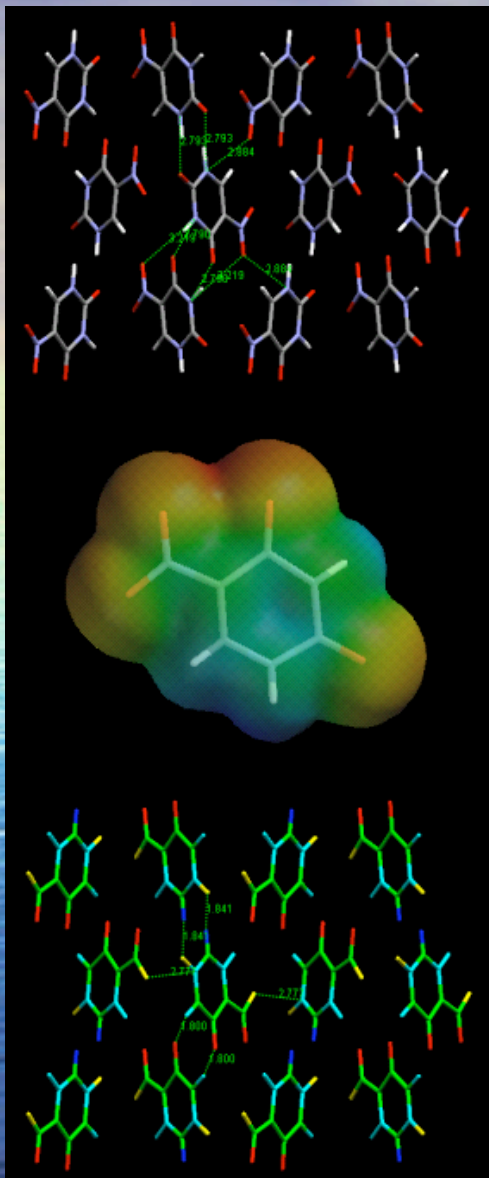
End - to - end connectivity

- Provide the smooth connection between the sources of data & information
- From literature to the laboratory bench and back via all stages of analysis and discussion
- Thus the need for a Data Grid or Grids
- All steps need to be Grid aware

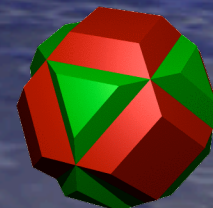


Generate information within & for the grid context



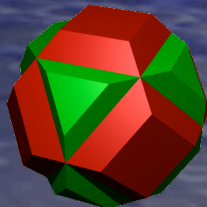


Variety of data



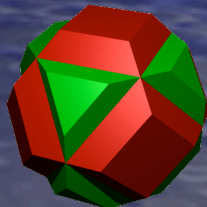
The Grid

- Grid is needed because
 - Complexity of data
 - Volume of data (real time data, images, video)
 - Scale of computation (analysis, simulation)
 - Complexity of process (automation)
 - Variable demands on computation
 - Provenance (audit trails, timestamps, process)



Dissemination & Publication

- A different approach is required to provide data to the community
- The grid provides the necessary medium
- What & How do we want to make available



Journals: Publication @ source

Database

Journal

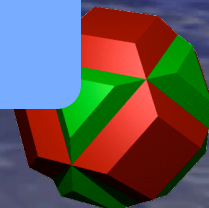
Paper

Materials

Laboratory Data

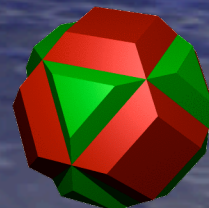
Multimedia

“Full” record



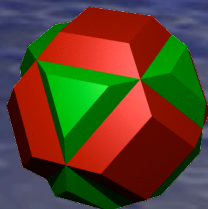
Data Trail

- Drill down through the analysis path
- Look at increasingly raw data
- Often large expansion in quantity and variety at each stage
- Need URIs for everything

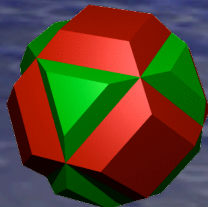
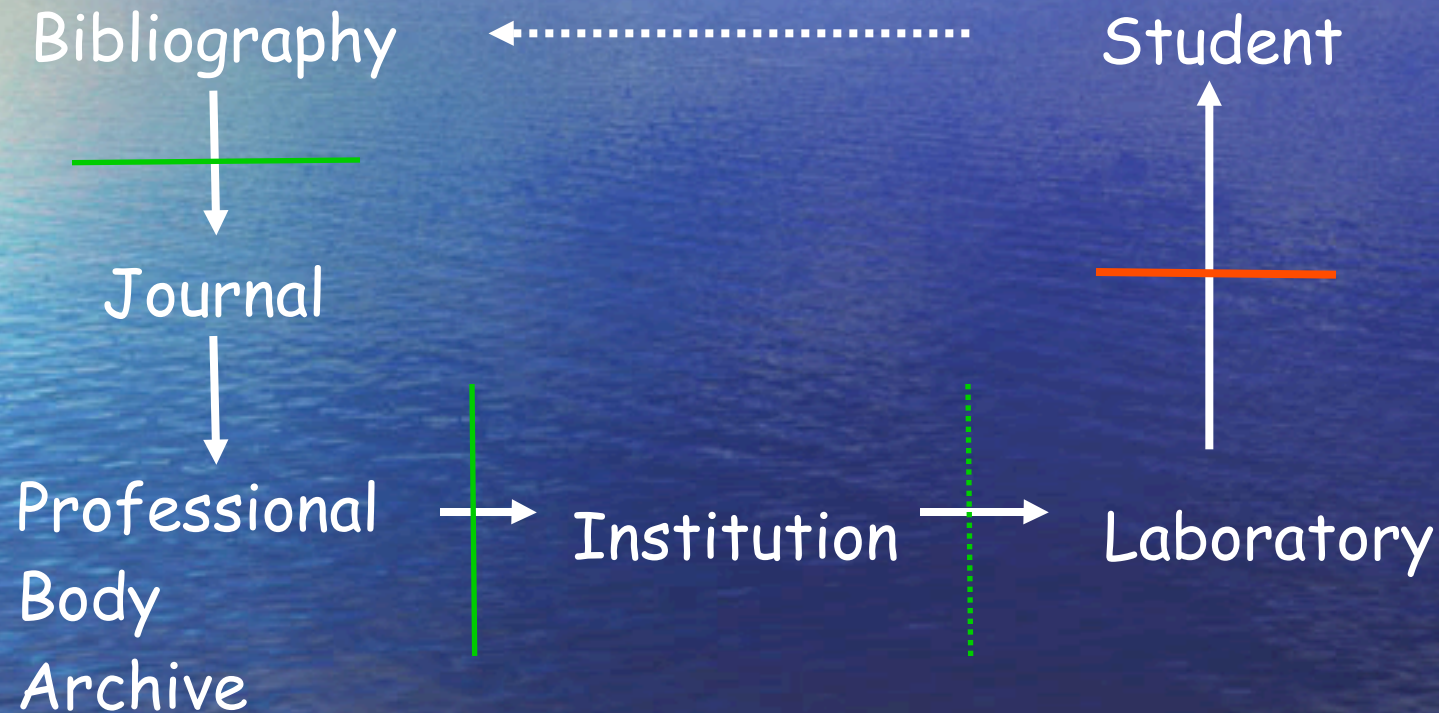


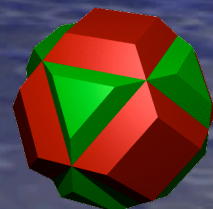
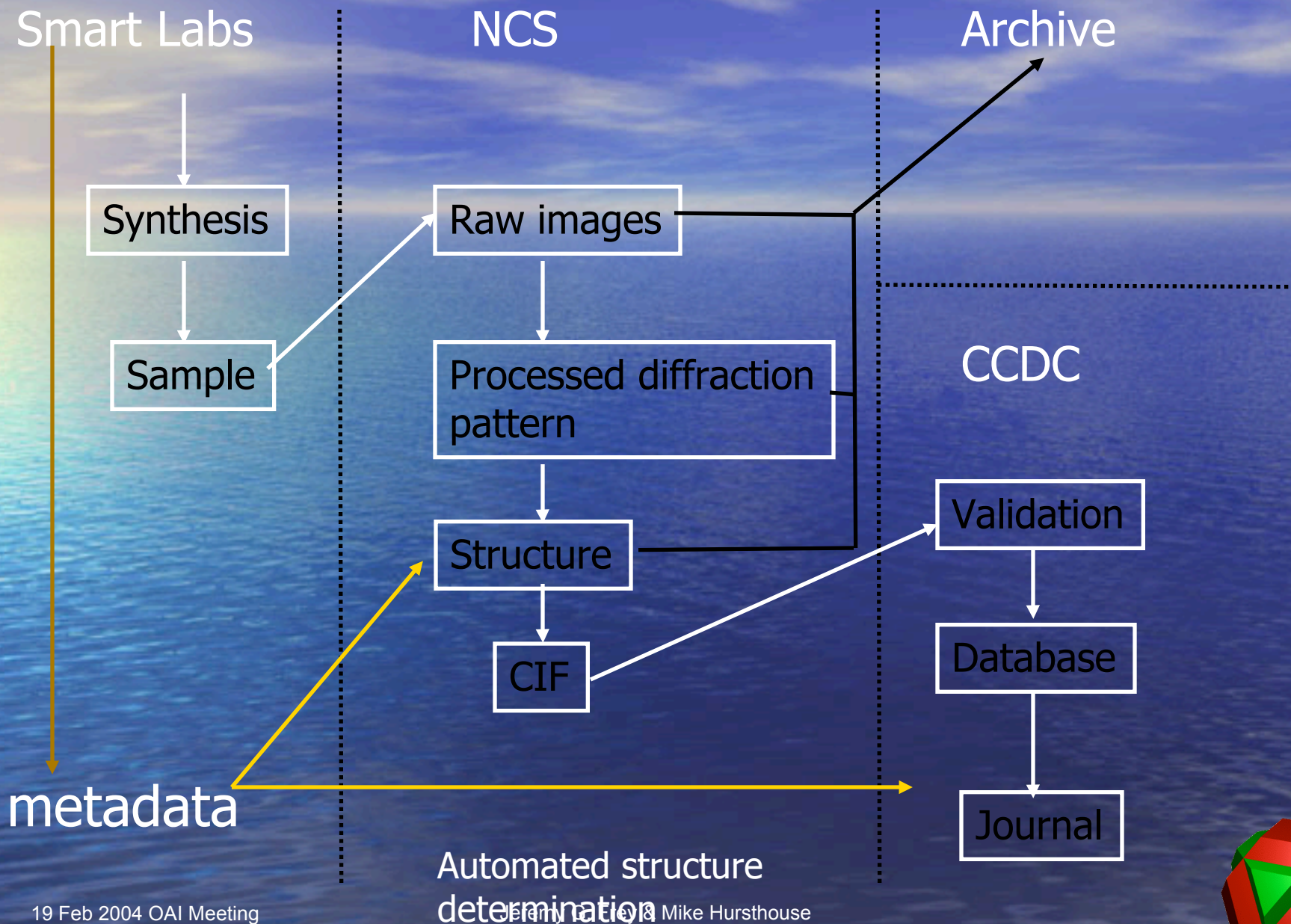
Publication@Source

- Must be able to track back to the original data
- Primary reason is to allow new analysis in the future by other researchers.
- In a university environment this may be viewed as a public responsibility in business environment ensuring maximum value from investment.
- Does have implications for provenance and even fraud!



Publication Chain







Chemical Crystallography: A Suitable Case for OA Therapy

Mike Hursthouse

Department of Chemistry and Combinatorial
Centre of Excellence,

EPSRC National Service for Crystallography

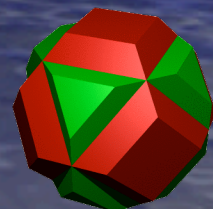
University of Southampton, UK



Comb-e-Chem

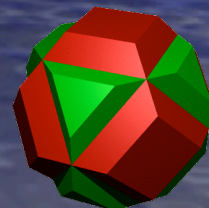
ChemCryst

- Characterisation technique for Chemical Structure.
- Use XRD.
- Provides high level of chem knowledge
- Structure – molecular or crystal
- Previously focussed on molecular structure – chemical props
- Now focus on crystal structure – physical props
- Change in interest facilitated by availability of database archive.
- However, woefully incomplete



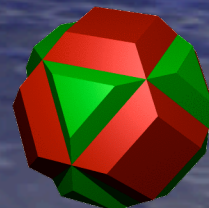
ChemCryst

- Database Archive – ca 300000 entries – all published structures
- >10M chemical compounds known
- Probably 1.5M structures known
- Why shortfall? Archaic publishing methods.
- Solution?



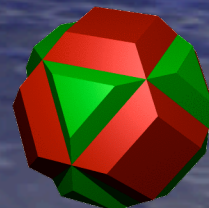
ChemCryst

- ChemCryst results New dissemination strategy
- E-Prints of “Structure Reports”
- Can be created automatically.
- Work can be validated automatically.
- All data (raw, processed, meta...) included.
- Hence bypass Journal sponsored “refereeing”
- Still need to decide on “publication” of “science”



e-Bank Project JISC project with UKOLN

- Link comb-e-chem and other semantic grid science projects to the e-print system at Southampton
- Provide dissemination and provenance



Changing the way we work

