Science Ajar?
From Open Access to Intelligently Accessible Data in Open Notebooks

Jeremy G. Frey
University of Southampton

A Chemist’s Digital Aura

http://mosaically.com/PhotoMosaic/2389771f-4bf3-49bf-bc8a-1810167096ff/
Reducing and Managing Uncertainty

Coherent diffraction Imaging for lens-less soft x-ray microscopy

Key is computational phase reconstruction, which using GPU, can be done in time comparable to the experimental acquisition time. GPU cluster will enable ensemble solution and thus estimate the uncertainty in the reconstructing, but would need on demand.

Air/Water slab MD simulation

Reducing Uncertainty

• What is already known?
• Who is doing what?
• How was a result obtained?
• Access to Data
• Integration of Data.

• Provenance & Planning

Computational power now available means that sufficient accuracy and low uncertainty can be obtained for a whole series of small and large water slabs allowing the surface contribution to the total energy to be reliably “observed” and entropy contributions to surface tension derived.
Intelligent (Open) Access to Data


05/12/2014

Beyond Open Access

Confusion over open-data rules

The Public Library of Science’s pioneering open-data mandate has prompted scientists to share more data online, but not everyone is complying with the regulations.

BY RUSSELL GAINSBURG

T
he outcome of the open-data rules is that scientists should share all data underlying their findings—something that has long been a goal of the non-profit’s biggest project, the Public Library of Science (PLoS), whose mission is to publish high-quality research articles that are freely accessible.

But in practice, there is an ongoing debate over whether the new data-sharing guidelines will be effective in ensuring that researchers publish their data.

Some scientists argue that the new rules are too strict, while others say the guidelines are too lenient.

“The research community needs to come together to make sure that the new data-sharing guidelines are effective,” said one scientist.

Others say that the new rules are a good step forward, but more work needs to be done to ensure that all researchers comply.

“The new data-sharing guidelines are a significant step forward,” said another scientist.

05/12/2014

Beyond Open Access

“In theory, there is no difference between theory and practice. But, in practice, there is.”

Unknown (possibly Yogi Berra or Manfred Eigen)

http://www.nature.com/news/confusion-over-publisher-a-pioneering-open-data-rules-1.15409

05/12/2014

Beyond Open Access

Gates Foundation announces world’s strongest policy on open access research

Gates Foundation announces world’s strongest policy on open access research

21 Nov 2014 | 18:00 GMT | Posted by Richard Van Noort | Category: Policy, Publishing

The Bill and Melinda Gates Foundation has announced the world’s strongest policy in support of open research and open data.

“We believe that open data and open research are essential tools for advancing health and development,” said the foundation’s co-founder, Bill Gates.

“We are committed to making our research and data open by default, so that scientists and policymakers can use them to make decisions that improve lives around the world.”

The new policy, which is effective immediately, requires all researchers funded by the foundation to make their research data open and accessible.

“This is a significant step forward,” said the foundation’s executive director, Ken Frazier.

05/12/2014

Beyond Open Access


05/12/2014

Beyond Open Access
Methods are as important as the data

Ocean observatory project hits rough water

Problems will data management challenge US effort to monitor sea in real-time.

By ELIZABETH WYNN

F

or months, researchers have watched the

deepwater observatory, a network of

underwater sensors, in the north Pacific

Ocean. When the project hit rough waters

last month, it was a reminder of the

challenges that come with monitoring

the oceans.

The Ocean Observatories Initiative (OOI)

is a partnership between the National

Science Foundation, the University of

Washington, and several other

institutions. The network of sensors and

platforms collects data on a variety of

oceanic processes, including

atmospheric conditions, ocean

chemistry, and marine biology.

Despite these challenges, the OOI

remains a valuable tool for studying

the oceans.

http://www.nature.com/news/ocean-observatory-project-hits-rough-water-1.16406

Once Upon

Journal Cortex launches Registered Reports

Print or PDF this page

Dr Chris Chambers and Professor Sergio Della Sala | Associate Editor and Editor-in-Chief of the Elsevier Journal Cortex

On May 1st, Cortex launched a new innovation in scientific publishing called a Registered Report. Unlike conventional publishing models, Registered Reports split the review process into two stages. Initially, experimentally motivated and procedures proposed are reviewed before data are collected. Then, if peer reviews are favourable, offer authors "in-principle acceptance" of their paper. This guarantees publication of their future results providing that they adhere precisely to their registered protocol. Once their experiment is complete, authors then resubmit their full manuscript for final editorial.

http://editorupdate.elsevier.com/short-communications/journal-cortex-launches-registered-reports

http://997waystobagreatspeaker.com/2010/12/how-to-tell-a-story

http://news.bbc.co.uk/1/hi/technology/8388485.stm
Skills we need are story telling skills NOT data mining skills

Clive Humby
DE/2014 Keynote

Faraday's laboratory notebooks are also remarkable in the amount of detail that they give about the design and setting up of experiments, interspersed with comments about their outcome and thoughts of a more philosophical kind. All are couched in plain language, with many vivid phrases of delightful spontaneity....

Peter Day, 'The Philosopher's Tree: A Selection of Michael Faraday's Writings'

Scientific Data Value Chain

- Does the Web currently/have the potential to disrupt or support the data value chain?

http://www.dreamingnewmexico.org/food/ff-local-foodshed
The internet has also changed how we communicate with one another — and what is known as ‘Web 2.0’ is here. Blogs, social networking sites, wikis and podcasts were alien concepts not all that long ago, but they are here to stay and some of them are significantly impacting the communication of science, and chemistry is no exception. Nature Publishing Group joined the chemical biosphere back in March 2006 when we launched ‘The Sceptical Chymist’ blog and ‘ChemPod’ — a chemistry-themed podcast if you hadn't guessed. To reflect the firm inroads that chemistry seems to have made into the blogging culture, Nature Chemistry will feature a monthly column called ‘Blogrol’ that will round up stories that have caught the attention of the chemistry bloggers. Our first topics include a citation controversy and a reminder that borne isn't banned!

Nature Chemistry

Web 2.0 & Semantic Web

USER GENERATED CONTENT

Accuracy


QUALITY?

DIGITAL RESEARCH NOTEBOOKS
Open Notebook Science

- Certainly not always the way to work!
  - IPR, Commercial, long term projects, recognition issues, etc

  It is not necessary to change. Survival is not mandatory. W. Edwards Deming

- But
  - Makes connection much easier if the data and processes are “Open”
  - Easy to export & access of “Linked-Data”
LabTrove Open Notebooks
Mat Todd's Malaria Project

LabTrove -> RSS -> Email / Twitter

Finding the “Wood” From the “Trees”

http://upload.wikimedia.org/wikipedia/commons/4/41/Looking_through_the_Trees_in_Spearywell_Wood_.jpg

Organized Data
- Metadata
- Linking
- Networks/Graphs
A big proportion of our users are not really adding metadata!

Chemistry blogs results

50 chemistry-related blogs – 10,436 ‘tags’

Categorizing ‘activity’ nouns

NASA results

53 NASA blogs – 1226 ‘tags’
**Metadata**

- User Created
- Metadata Sentence
  - Attempt to get verbs (processes) as well as nouns (things)

```
http://en.wikipedia.org/wiki/Toki_Pona
```

---

**Smart Tea Project**

**P2 - Plans & Provenance**

```
Chemistry and Tea: An Interesting History
A Nice cup of Tea
- George Orwell

Chemists seek perfect cup of tea - for Orwell
- London (Reuters)

How to make the perfect cup of tea (off downed)
- The Royal Society of Chemistry provide the definitive recipe for a cup of tea
```

---

**Graph/Network provides intuitive navigation**

---

Research student thesis writing often involves dealing with the consequences of having delayed the (huge) organization effort to the point of writing up.
“The Laboratory”

- A laboratory is an integration of people and equipment
- A complex socio-technical system
- Disrupted or integrated by digital infrastructure?
- How to support (a) people to do what they are best at doing and (b) computers to undertake what they can do best and (c) ensure that they can work together.

“Engrooved*”

* concept highlighted in teaching by Paul Trowler
- [http://www.brad.ac.uk/sustainable-universities/media/sustainableuniversities/Keynote-Prof-Paul-Trowler.pdf](http://www.brad.ac.uk/sustainable-universities/media/sustainableuniversities/Keynote-Prof-Paul-Trowler.pdf)

Bringing dissemination in to the lab... use and re-use of data... the digital research notebook

DISSEMINATION IS PART OF THE RESEARCH
The System
adapted from Atul Gawande 2014 Reith Lectures

• “Primitive idea”
  – You ought to do x.
• Medieval approach
  – You must do x! You must follow the standards
  – Will will take away your grant....!
• Modern world(?)
  – Becomes the norm, make it easy to do
  – Checklists
  – Feedback loops

We must speed up the knowledge discovery process

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

Re-imagining Scientific Research

A “Social Machine” of experienced researchers, technology & information

LABORATORIES OF THE FUTURE
Everything is Awesome,
Everything is cool when you're part of a team
The Art of Asking the Right Question

He who asks a question is a fool for five minutes;
he who does not ask a question remains a fool forever.

— Chinese Proverb

Why is asking a good question on Quora (20 credits per follower) more rewarding than writing a good answer (10 credits per upvote)?

http://www.quora.com/Why-is-asking-a-good-question-on-Quora-20-credits-per-follower-more-rewarding-than-writing-a-good-answer-10-credits-per-upvote

Acknowledgements

• My research group
  – Chemists & Computer Scientists
  – UG project students, PG, and Post-Docs
  – Visitors

• My colleagues
  – Chemistry, Physics & ORC, Computing, Engineering,
    – Statistics/ Mathematics

• Funding
  – University of Southampton, Chemistry
  – EPSRC, RCUK, JISC/Jisc, IBM, Microsoft, TSB, Lonza

05/12/2014
Beyond Open Access

Thank you for listening

Trust me Mort - no electronic communications superhighway, no matter how vast and sophisticated, will ever replace the art of the schmooze