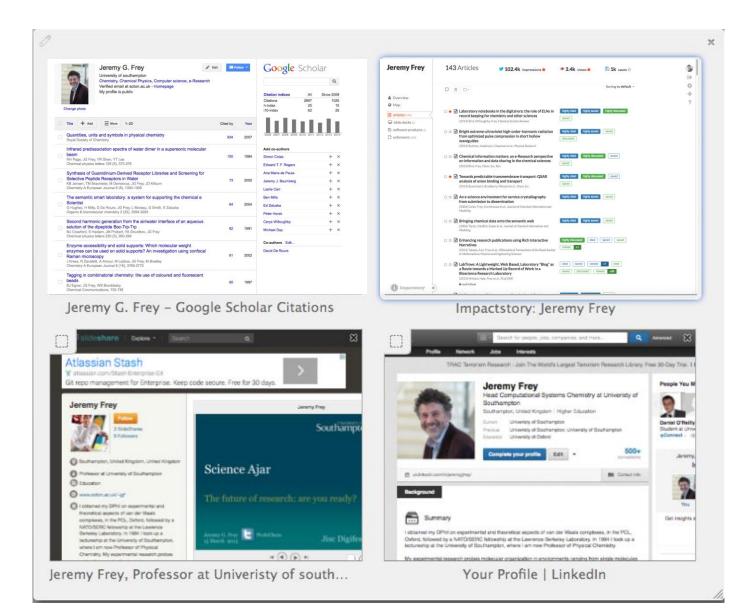


Science Ajar? From Open Access to Intelligently Accessible Data in Open Notebooks

Jeremy G. Frey University of Southampton

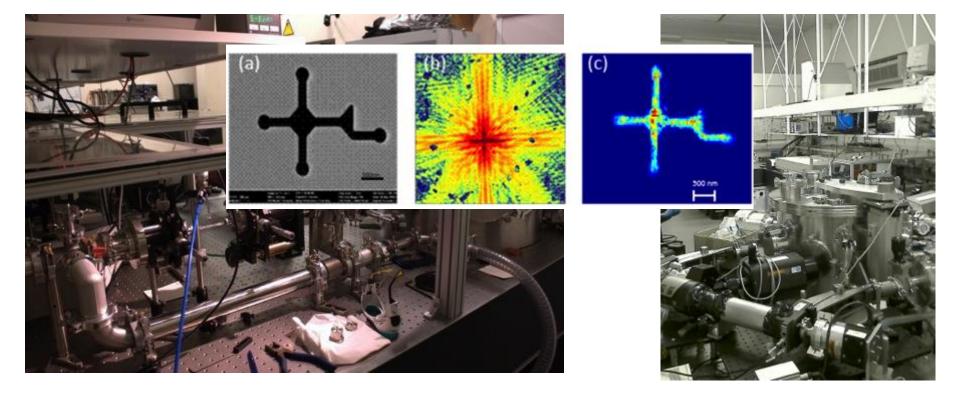
A Chemist's Digital Aura



A Chemist's Digital Aura



http://mosaically.com/PhotoMosaic/3289771f-0bf3-49bf-bc8a-1810167096ff



UltraFast Xray Group

<< Next Post

Reconstructing the new data

3rd August 2011 @ 22:52

Just a quick post to show some pictures from the reconstructions of the latest data and update everyone on progress.

Looking at data taken on 2nd August - the old 2um sample taken with a single wavelength. The diffraction patterns look v clean, with good detail right out to the edge - haven't worked out the resolution yet.

Tried to get Ben's CDI routine working, but it broke badly, and I couldn't fix i immediately, so i N went back to my old routines (from 2 years agol can't believe it's been that long) and got 'hio6.m' working. This is running on Boxer under Windows, but not using anything clever.



LISSE_intertionLasal This one is a very plain HIO reconstruction, 1000 iterations after binning the data to 512x512. Shrinkwrapping every 100 0 think - notes at work right nowi. You can see that it reconstructs beautifully, and this happens every time - no need for multiple phases to get it to work.

(Once I'd remembered to background-subtract properly, of course - before that it all failed May 2011 (23) April 2011 (17)

	ITY OF	
outi	nam	ptor

Previous Post >>

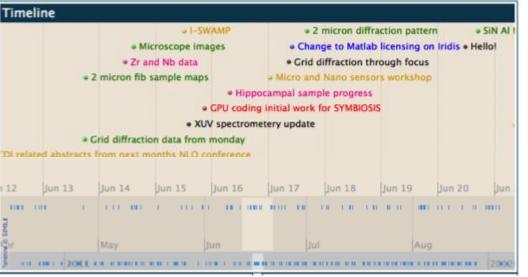
Search This Post Permalink

URI Label Revisions Export:

XML (With Files) PNG Image This Blog

New Post Blog Settings Timeline View Exhibit View Export Blog Archives November 2011 (18) October 2011 (40) September 2011 (38)

September 2011 (38) August 2011 (40) July 2011 (31) June 2011 (35) May 2011 (23)



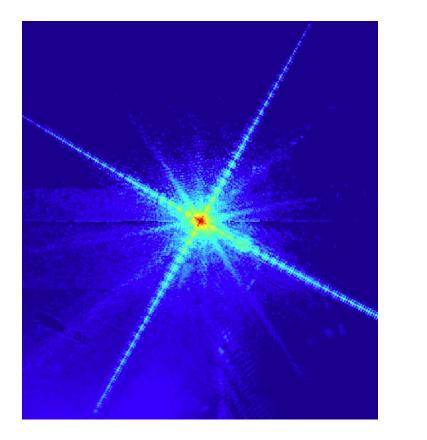
The future of laboratory data

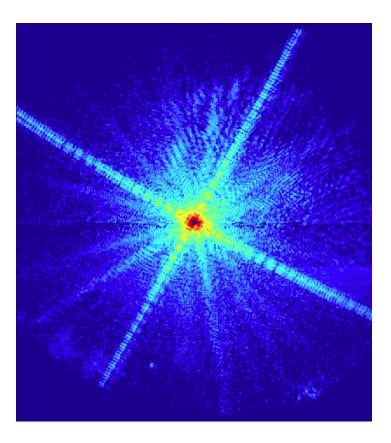
http://commons.wikimedia.org/wiki/File:ISS-30_Ice_floes_along_the_Kamchatka_coastline. jpg



Reducing and Managing Uncertainty

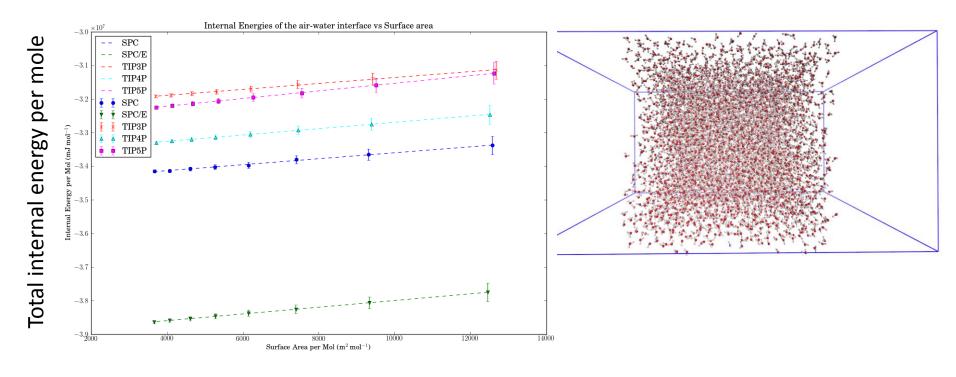
Coherent diffraction Imaging for lensless soft x-ray microscopy





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



Surface area per molecule

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.

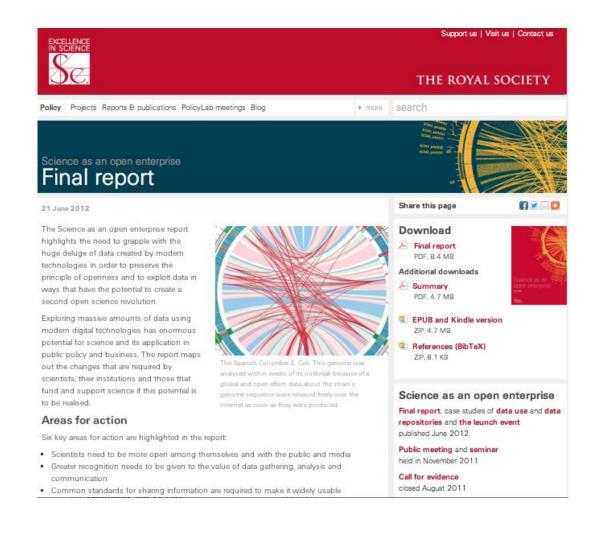
Reducing Uncertainty

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

• Provenance & Planning



Intelligent (Open) Access to Data



https://royalsociety.org/polic y/projects/science-publicenterprise/Report/



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 regulation.

BY RICHARD VAN NOORDEN

The mantra of the nascent open-data movement — that scientists should share online all data underlying their findings — sounds simple. But it can be tough to achieve in practice. An informal audit of one of the movement's biggest proponents, the Public Library of Science (PLOS), shows that not everyone is complying with the publisher's pioneering open-data mandate, and hints at the challenges that journals can face in enforcing open-data goals.

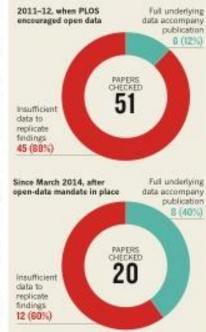
The idea that the progress of research will be accelerated if others can easily and freely build on data sets is gaining currency. Last week the Bill & Melinda Gates Foundation in Seattle, Washington, announced that it would demand open data of the researchers it funds.

But whereas some research communities, such as geneticists and crystallographers, have long-established norms of open data, most funders and publishers (including *Nature*), mindful of researcher autonomy, merely exhort scientists to make their data open. Many surveys have found that scientists are worried about being scooped on future projects, or argue that they have signed agreements not to share their data.

So it was a step-change when in March PLOS made it a requirement that authors who publish in its journals share online all the data necessary to reproduce their studies. It was not

FREE THE DATA

In 2014, open-access publisher PLOS introduced a requirement that authors who publish in its journals make their underlying data freely available online. An informal audit of one type of population genetics study in one journal, PLoS ONE, shows that not everyone is complying — but the mandate is still a boon for the open-data movement.



job of external peer reviewers to check whether appropriate underlying data were available, he says. And although PLoS ONE was grateful for their help, "there is a learning curve here for all involved to understand what the data-sharing standards are for all disciplines and data types". He adds that once someone complains, the journal has a system for investigating papers that do not comply with its open-data requirement.

The research teams in question, nine of whom responded to the Nature news team's request for an explanation, provide an insight into why data sharing does not always happen in the first place (full data are available at go. nature.com/8ggob6). Some had forgotten to upload their data and promptly rectified the fault. Four teams said that the journal's editors and referees had never asked them to share the underlying genetics data, suggesting confusion over what the policy means.

Others aired wider — and common — objections to online data sharing. Even though they had chosen to publish with PLOS, some authors said that they wanted to hold back their data for future studies, or did not want to share the raw data unless they knew future users' intent.

Steve Simpson at the University of Exeter, UK, who reported work on Omani clownfish, said he was happy to share raw data privately with potential collaborators, but not to upload results from 400 individual fish that had taken great effort to collect. "The study is described so that it could be replicated by another expe-

http://www.nature.com/news/confusion-over-publisher-s-pioneeringopen-data-rules-1.16409



newsblog

Nature brings you breaking news from the world of science

News & Comment

News Blog > Post

Previous post

Energy outlook sees continuing dominance of fossil fuels



NATURE NEWS BLOG

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

21 Nov 2014 | 18:39 GMT | Posted by Richard Van Noorden | Category: Policy, Publishing

The Bill and Melinda Gates Foundation has announced the world's strongest policy in support of open research and open data. If strictly enforced, it would prevent Gates-funded researchers from publishing in well-known journals such as *Nature* and *Science*.

On 20 November, the medical charity, based in Seattle, Washington, <u>announced</u> that from January 2015, researchers it funds must make open their resulting papers and underlying data sets immediately upon publication — and must make that research available for commercial reuse. "We believe that published research resulting from our funding should be promptly and broadly disseminated," the foundation states. It says it will pay the necessary publication fees (which often amount to thousands of dollars per article).

http://blogs.nature.com/news/2014/11/gates-foundation-announcesworlds-strongest-policy-on-open-access-research.html

Beyound Open Access



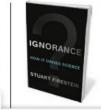
PHILOSOPHY

What we don't know

Michael Shermer enjoys a reminder that cutting-edge research is a step into the unknown.

t a press conference in 2002, Donald Rumsfeld, then US secretary of defence, used epistemology to explain US foreign entanglements and their unintended consequences. "There are known knowns; there are things we know we know. We also know there are known

Click for Article Access Options



Firestein introduces the concept of ignorance by contrasting the public's perception of science — as a systematic process — with a scientists'

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

Unknown (possibly Yogi Berra or Manfred Eigen)

Methods are as important as the data



http://news.bbc.co.uk/1/hi/technology /8388485.stm



B May 2013 © 6 Comments 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, experimental methods and proposed analyses are pre-registered and reviewed before data are collected. Then, if peer reviews are favourable, we 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 consideration.

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



MARINE SCIENCE

Ocean observatory project hits rough water

Problems with data management challenge US effort to monitor seas in real time.

BY ALEXANDRA WITZE

From the waters off western North America to the seas surrounding Greenland, US oceanographers have nearly finished deploying hundreds of sensors, moorings, gliders and other equipment that make up an ambitious US\$386-million effort to establish the world's biggest interactive portal to the oceans. Through the Ocean Observatories Initiative (OOI), set for completion by May 2015, anyone could watch the climate changing in the North Pacific or an underwater volcano erupting in real time.

But even as the final instruments splash

into the sea, the project has hit a snag. After spending \$37 million to develop state-of-theart software to manage live data, the OOI has terminated that contract — shifting responsibility for 'cyberinfrastructure' from the University of California, San Diego (UCSD), to a group based at Rutgers University in New Brunswick, New Jersey.

The shift is meant to help the OOI finish construction on time and on budget by next spring, and is likely to delay the start of data streaming by only a few months. Still, it is making some oceanographers anxious, given the project's broad scope. Data from its observatories are likely to dominate US oceanography for the next 25 years - the OOI's planned lifetime.

The project, funded by the US National Science Foundation, is intended to be for the oceans what earthquake and volcano monitors are for the land: a way to allow real-time glimpses into the currents, marine life and chemistry in the deep sea. It will scrutinize a few select patches of ocean in exhaustive detail, using instruments attached to cables and on free-sailing gliders to measure temperature, salinity, chemistry and other key oceanographic parameters. The network's scattered locations include a cabled sea-floor observatory off the coast of Oregon; sets of instrumented moorings off the US east and west coasts, and

474 | NATURE | VOL 515 | 27 NOVEMBER 2014

© 2014 Macmillars Publishers Limited: All rights reserved

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





http://997waystobeagreatspeaker.com/2010/1 2/how-to-tell-a-story/ Skills we need are story telling skills not data mining skills

Clive Humby DE/2014 Keynote



"We have lots of information technology. We just don't have any information."

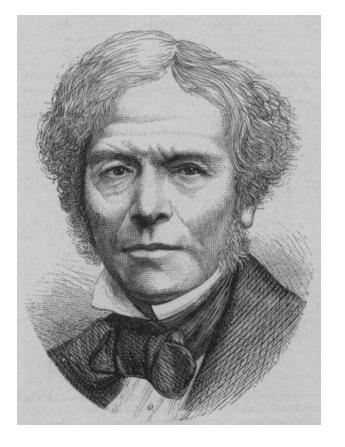
What is the story? What is the how and the why?

THE NARRATIVE

ON BEING A SCIENTIST

A GUIDE TO RESPONSIBLE CONDUCT IN RESEARCH

THIRD EDITION



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....

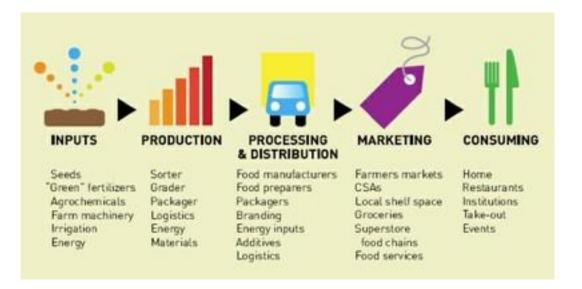
Peter Day, 'The Philosopher's Tree: A Selection of Michael Faraday's Writings'

WWW

THE INTERNET AND THE WEB

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 blogosphere 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 'Blogroll' 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 boron isn't borino!

Nature Chemistry

Web 2.0 & Semantic Web

USER GENERATED CONTENT

QUALITY?



http://www.opsrules.com/supply-chain-optimizationblog/bid/282916/Why-You-Shouldn-t-Waste-Effort-Improving-Forecast-Accuracy Beyound Open Access

DIGITAL RESEARCH NOTEBOOKS

The Planning Process Materials procedure Into the Lab! Into the Lab! dd Experim. MoreTea: Tablet Interface and RDF World

Images and sketch commenting Mutagenesis of plasmid p042 via Taq and Transformation of plasmid JRH4712/66 into preserving the record BW25141 by electroporation Test digestions to check the activity of a LabTrove ons were set up according to the following protocol. 37 "C briefly ons were set up according to the following protocol. 11 'S constant and abmose plates and SOC medium were una to the set of the set of the over plates were enround with Y constant. vicilin arabinose plates and SOC medium were warmed to 27 °C briefly arabinose plates and SOC medium were warmed to 21 X-giu and LBI arabinose plates were spread with X-giu (80 µL, 1: X-giu and LBI arb continue warming Lab Book Ref: jrh4712-78 batches of EcoRI and NCOl Lab Book Re 1814712/66, p042, and electroppoator Book Ret 1712-78 (1) -712-80 (1) 712-81 112.83 (1, 4712-82 (1. 4712-841np arabinose plates and 4712-85 (1) 712-88 (1) verences and the transformant transformation and the transformation transformation and the transformation and transformation and the transformation and the transformation and transformation 1712-69 12 r for one hour with sh 4712-86 (1) 4712-87 (1) 20 with LB and 100 P 712-9001 112-86 (). • run on program MUTACBG for 30 Ovcles. 5 µL product over a program MUTACBG for 30 Ovcles. 5 µL product inst sometic had failed. ·C overnight. The reactions were run on program MUTAGBG for 30 cycles. 5 µL product failed, s run on a 10% normal agarose a na/vical gel. Tag mutagenesis had failed, s run on a 10% normal agarose a na/vical gel. and was correct. product -712-903 112.74 12 product *CoTaq=1 µL stock + 3 µL ^{water.} 712-760 n on a 1% normal agarose analytical gel. Taq mutage. and +ve crr had worked. -ve control was correct. Sample Pares ifer Hale | Beta-glucuron bath at 37 °C for 3 hours. LabTrove: Easy Comments (2)

Highly inter-linked report

🖁 Digital and Physical

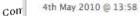
History <u>B</u>ookmarks Iools <u>H</u>elp Luke Marazzi's Y3 Chemistry Project

Previous Post >>

🕲 http://chemtools.chem.soton.ac.uk - barcode_print.php (application/p...

🕞 📋 🌟 | 🚑 • | 💠 💠 1 / 1 | 🔊 🖲 180% • |

TLT660/4b Hot Stage Video



🕙 Sortase Cloning - Mozilla Firefox

Edit File

> This wa≤ see



chem.soton.ac.uk/P

| View Source | Samples | Comments (0)

Add comment to Post

Title*

Re: TLT660/4b Hot Stage Video

Text*

Done

👭 start



This Post

Search

0

Permalink URI URI Label Revisions Export: XML

This Blog

New Post **Timeline View** Export Blog

Archives

April 2010 (1) March 2010 (1) March 2010 (2)

Authors

Sections

Appendix (1) Samples (3)

Tools Show/Hide OR Code Show/Hide Keys

QR Code



(suthampton



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" abbrove Calories today via



Public Blog Post: Synthesis of 3-(4morpholinothieno[3,2-d]pyrimidin-6yl)benzenesulfonamid...



determine the water content of thieno[3,2d]pyrimidi... <u>http://t.co/GnTcN09</u>l #<u>malaria</u> #<u>drugdesign</u>

• labtrove, [+] Fri 17 Aug 05:33 via twitterfeed



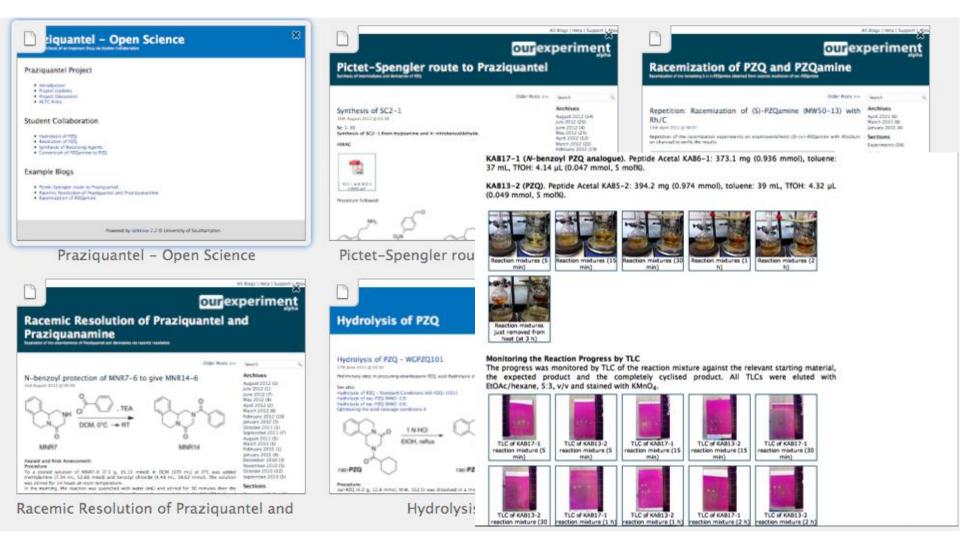
(# TVIMed) is out! <u>http://t.co/UNtoGQ</u> Top stories today via @<u>BiblioBuster</u> @<u>labtrove</u> @<u>a100ciacierta</u>

• RonanKelly13, [+] Thu 16 Aug 11:27 via Paper.li





LabTrove Open Notebooks Mat Todd's Malaria Project



LabTrove -> RSS -> Email / Twitter

IFTTT			My Recipes	Browse	Channels	jgfrey -	
My Recipe	S						
Personal	Shared	Favorites	2 Personal Recipes, 2 turned or				

Personal Recipes are a combination of a Trigger and an Action from your active Channels.

Create a Recipe

Filter -





If new feed item from http://xray.orc.soton.ac.uk/feeds /xray_group

/uid/2e575a007d7c25e2e54e672d1954b842?withcom ments, then send me an email at

created less than a minute ago never triggered

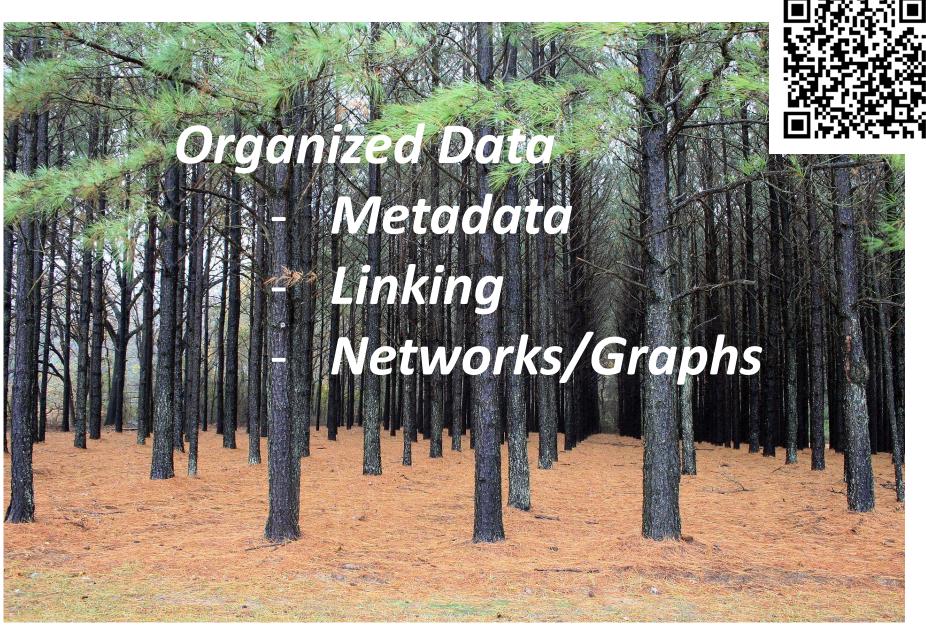




created February 25, 2014 last triggered April 02, 2014 triggered 10 times



http://upload.wikimedia.org/wikipedia/commons/4/41/Looking_through_the_Trees_in_S pearywelf_Wood_-_geograph.org.uk_-_424274.jpg

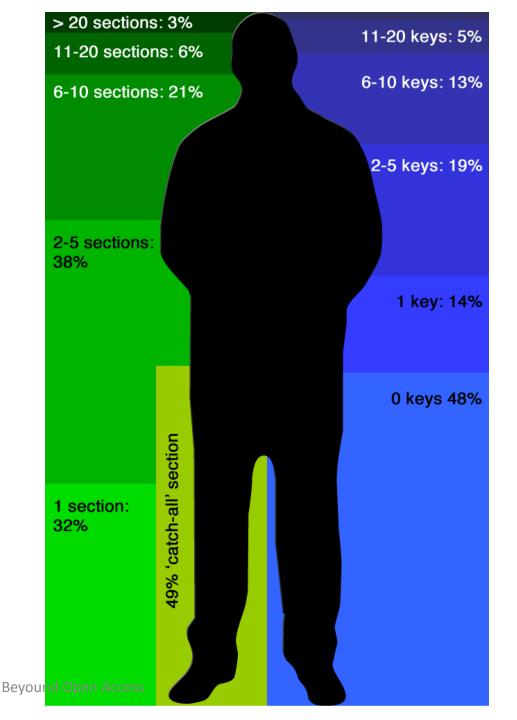


http://en.wikipedia.org/wiki/Plantation#mediaviewer/File:Pinus_taeda_plantation.jpg

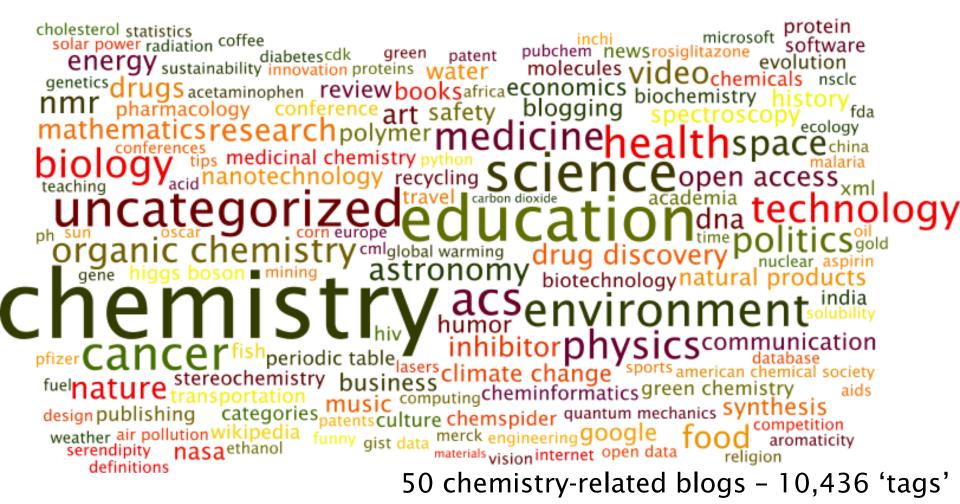
05/12/2014

Beyound Open Access

Abig proportion of our users are not really adding metadata!



Chemistry blogs results



Beyound Open Access

Categorizing 'activity' nouns Tetles Lo LEZ Trav

Mai

NE

Distan La 152

Zerti

Tellins Lot

maters 544 TEXT

Taraba

Units!

LSZD Tatal

57.

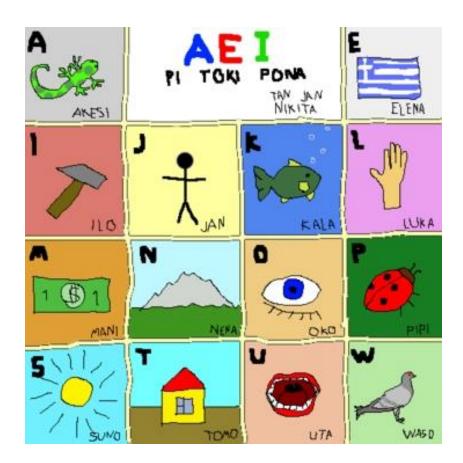
NASA results



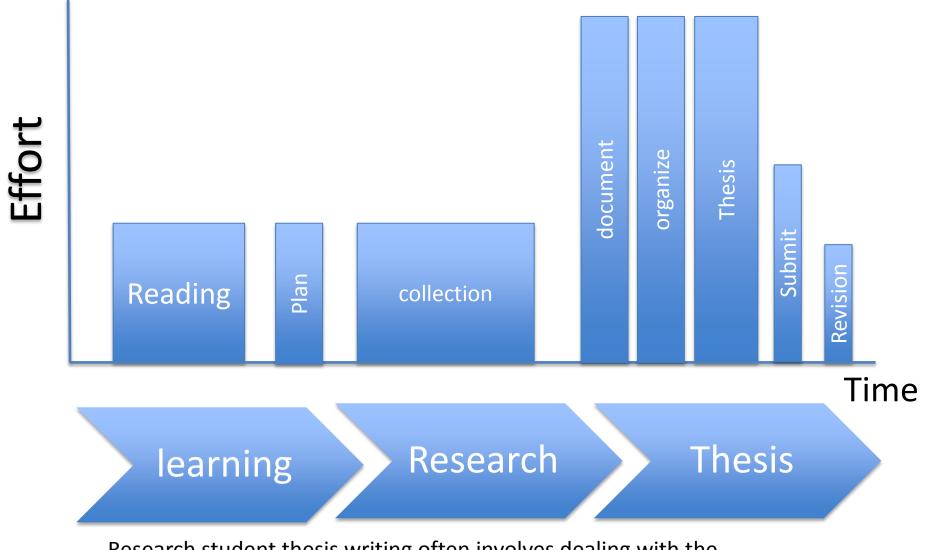
Beyound Open Access 53 NASA blogs – 1226 'tags'

Metadata

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



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



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

Smart Tea Project

the Small Tea Project and its, like, this is the standard of a new of the standard of a new of the standard of a new of the standard of the st

Chemistry and Tea: An Interesting History

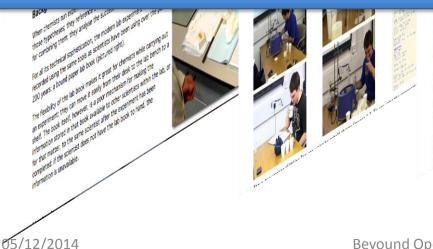
A Nice cup of Tea

George Orwell

<u>Chemists seek perfect cup</u> of tea – for Orwell – London (Reuters)

How to make the perfect <u>cup of tea</u> (pdf download) – The Royal Society of Chemistry provide the "definitive recipe for a cup of tea"

P2 - Plans & Provenance

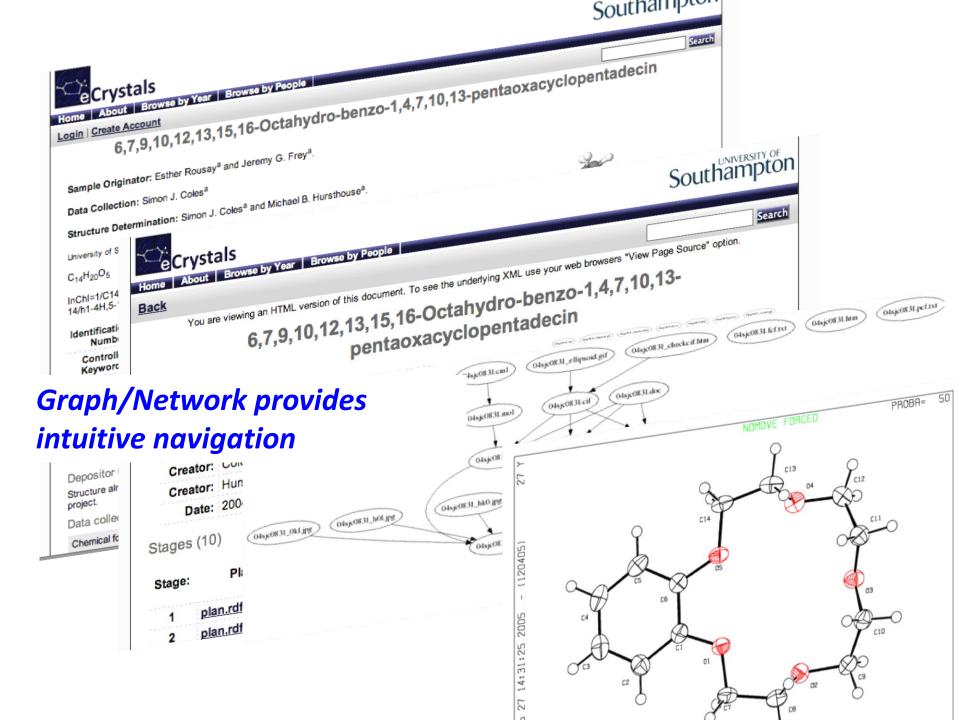


enter standard no.: BS6008:1980 (6 pages, £22.00) Guardian story (and abbreviated version) of this BSI prize winning document.

Counterfeit tea (or smouch)

 The BBC (and Michael Quinion for the explanation of smouch)

Beyound Open Access



VALIDATION

Publication@Source

"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.



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

DISSEMINATION IS PART OF THE RESEARCH

"Engrooved*"

- * concept highlighted in teaching by Paul Trowler
 - http://www.brad.ac.uk/s ustainableuniversities/media/susta inableuniversities/Keyno te-Prof-Paul-Trowler.pdf



http://www.todayandtomorrow.net/201 0/02/22/quarter-mile-groove/

The System

adapted from Atul Gawande 2014 Reith Lectures

- "Primitive idea"
 - You ought to do x.
- Medieval approach
 - You <u>must</u> 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



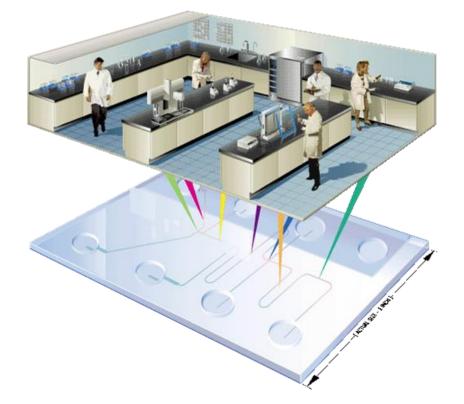
All I am saying is that now is the time to develop the technology to deflect an asteroid

© The New Yorker collection. All rights reserved from The New Yorker Book of Technology Carto



Re-imagining Scientific Research

A "Social Machine" of experienced researchers, technology & information http://lab-on-chip.gene-quantification.info/



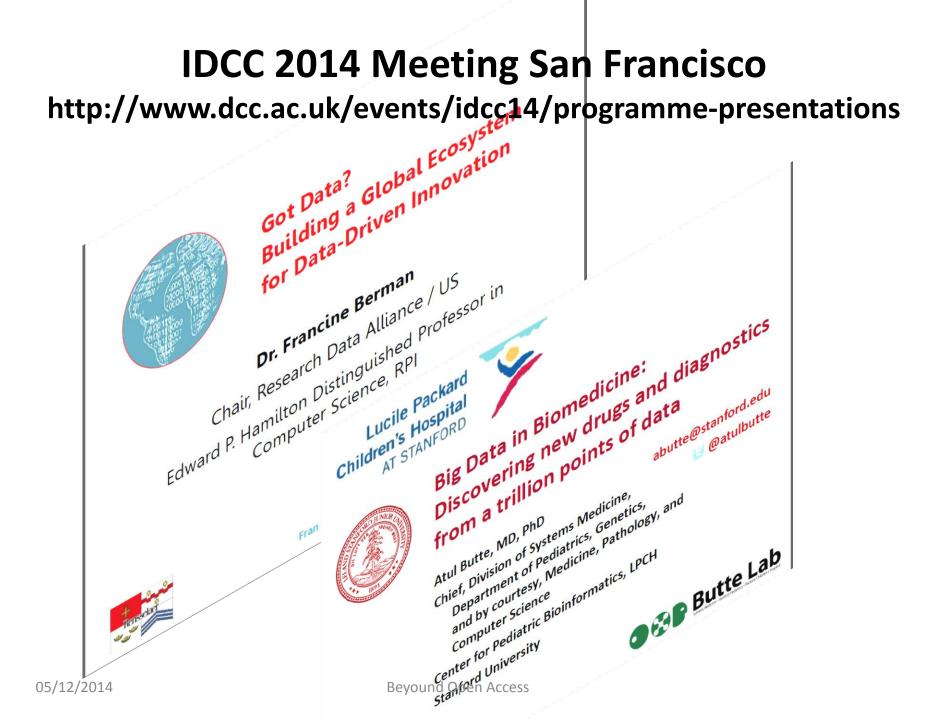
LABORATORIES OF THE FUTURE

The Lego Movie



Everything is Awesome, Everything is cool when you're part of a team







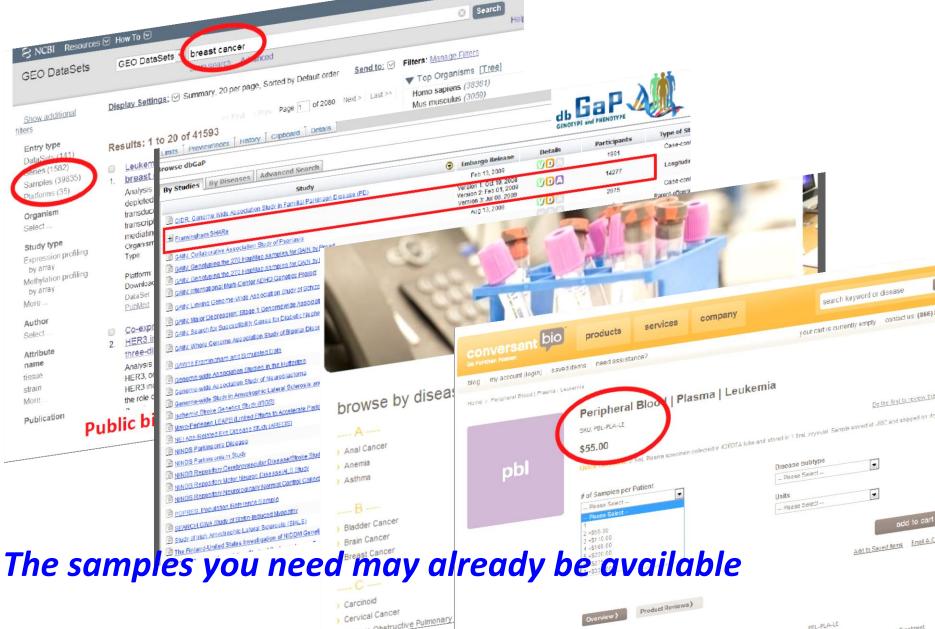
Beyound Open Access

Additional Information SKU

Treatment Status

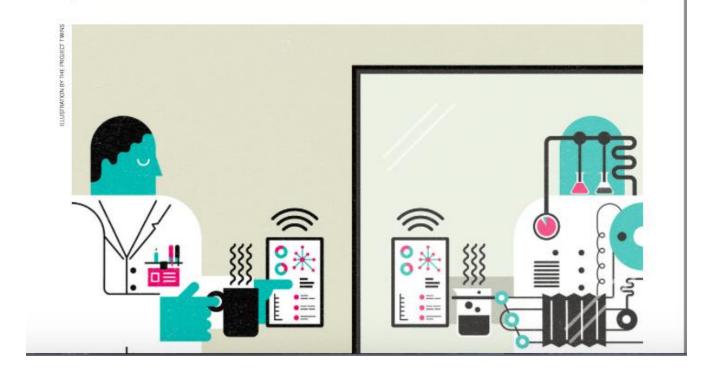
Any Treatment, Pre Treatment, Post Treatment PEL-PLA-LE Active Treatment, Pre-Treatment, Post-Treatment, Active Treatment, Recurrent/Refractory Disease, Remission

Chronic Obstructive Pulmonary



THE Automated Lab

Start-up firms say robotics and software that autonomously record every detail of an experiment can transform the efficiency and reliability of research.



http://www.nature.com/polopoly_fs/1.16429!/menu/main/topColumns/topLeftColumn/pdf/516131a.pdf

Beyound Open Access

The Art of Asking the Right Question

Are You Asking The Right Question?

+ Comment Now + Follow Comments

Forty years ago, someone asked a profound question that fundamentally changed how we communicate with each other every day. At the time, I was nine years old living in South Africa. When our family visited my grandfather's farm for the holidays, the telephone we used was a "party line", where we would listen for the pattern of rings when the telephone rang to determine if the call was for us or for the next farm over. Hard to believe now—it does not seem that long ago.

At the same time, Marty, a young engineer at Motorola, was given a new assignment. He was asked to lead a team on a project that showed great promise—the next generation of a car radiotelephone. Marty accepted the challenge. However, instead of jumping in, he stepped back and paused, which led him to ask himself a very insightful question.

"Why is it that when we want to call and talk to a person, we have to call a place?" That nagging, insightful question changed the entire trajectory of his work, as he refocused his team's attention on untethering a person from a place (including a car).



Want to Lead? Learn to Ask the Right Questions

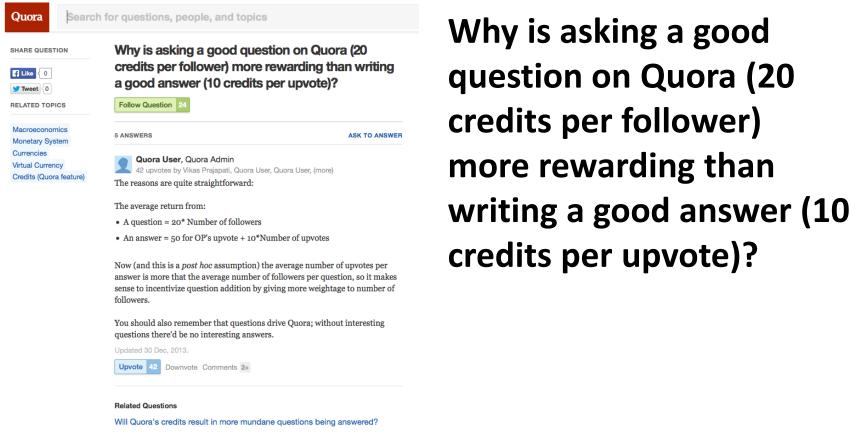
by BIG THINK EDITORS MAY 22, 2014, 12:00 AM



Have you ever walked away from a sales meeting wondering what you could have said differently to a potential client? Or do you feel as though you're only wasting your breath trying to convince your employees to believe in a new initiative or direction? Are you tired of telling your children to clean up their rooms? Convincing others can't come from you—it has to come from them. How does this happen? Learn to ask the right questions.

Daniel Pink is a management expert and the author of five bestselling business books, including *Drive*. He stopped by *Big Think*'s studio to talk about his latest book *To Sell is Human: The Surprising Truth About Moving Others*. *He who asks a question is a fool for five minutes; he who does not ask a question remains a fool forever.*

-Chinese Proverb



How many credits do you get for upvotes on answers? And does the author of the question get any credits?

What is the goal of Quora Credits?

http://www.quora.com/Why-is-asking-a-good-question-on-Quora-20-credits-per-follower-more-rewarding-thanwriting-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



Thank you for listening

collection. All rights reserved ter Book of Technology Cartoon

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