



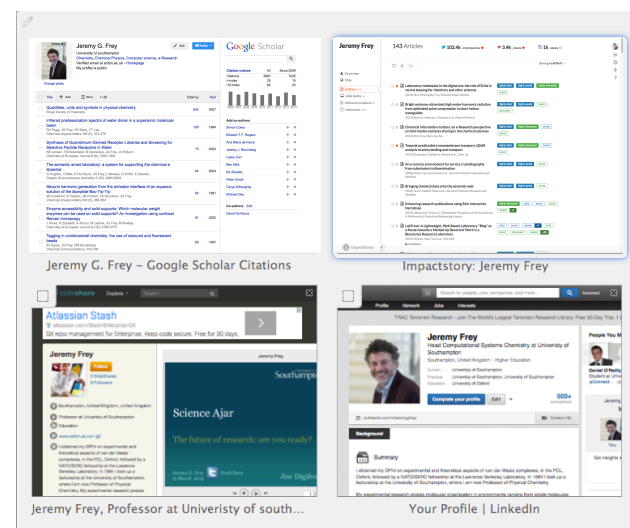
Science Ajar? From Open Access to Intelligently Accessible Data in Open Notebooks

Jeremy G. Frey
University of Southampton

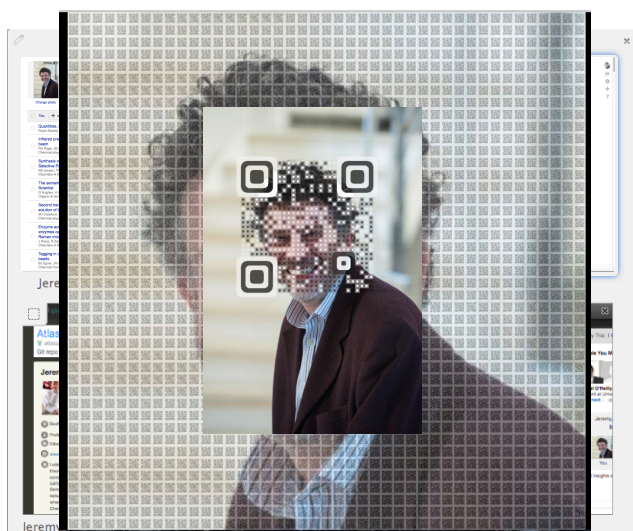
05/12/2014

Beyond Open Access

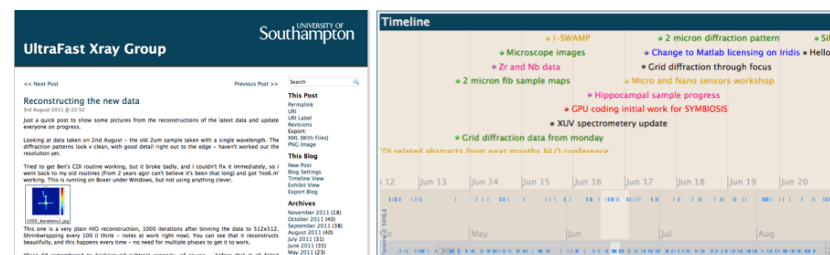
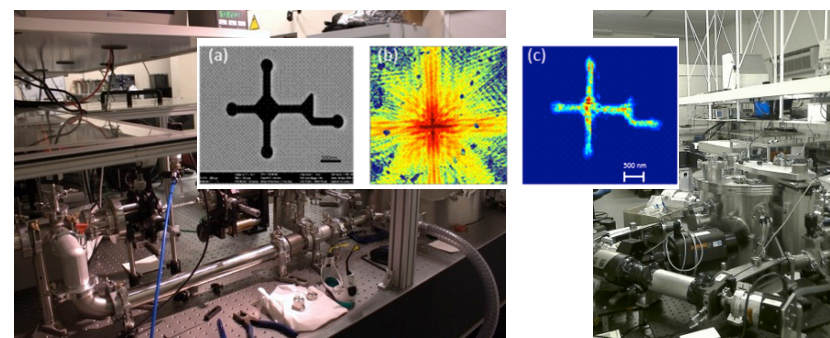
A Chemist's Digital Aura



A Chemist's Digital Aura



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



05/12/14

The future of laboratory data

4

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

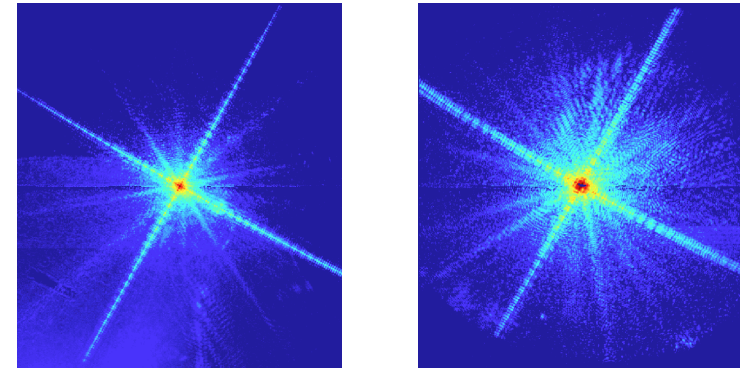


Reducing and Managing Uncertainty

05/12/2014

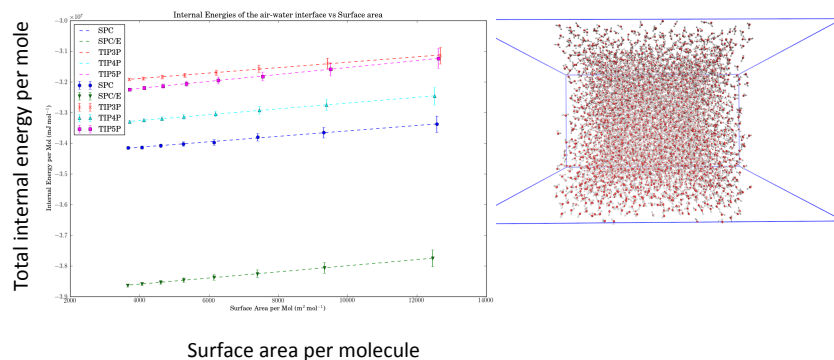
Beyond Open Access

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.

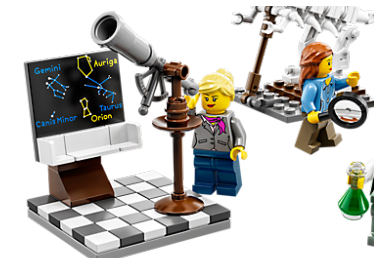
Air/Water slab MD simulation



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/policy/projects/science-public-enterprise/Report/>



05/12/2014

Beyond Open Access



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

05/12/2014

Beyond Open Access

PUBLISHING

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

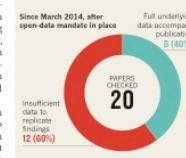
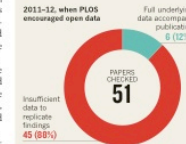
The mantra of the nascent open-data movement — that scientists should share online all data underlying their findings — sounds simple. But it can be tough to achieve in practice. An informal audit of one of the movement's biggest proponents, the Public Library of Science (PLOS), shows that not everyone 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 teams 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/9gob66). 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 exper-

<http://www.nature.com/news/confusion-over-publisher-s-pioneering-open-data-rules-1.16409>



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

Unknown (possibly Yogi Berra or Manfred Eigen)

05/12/2014

Beyond Open Access

Methods are as important as the data



05/12/2014

Beyond Open Access

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



05/12/2014

Beyond Open Access

<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 WITZ

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-the-art 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 Macmillan Publishers Limited. All rights reserved.

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

05/12/2014

Beyond Open Access



<http://997waystobeagreatspeaker.com/2010/12/how-to-tell-a-story/>

05/12/2014

Beyond Open Access

*Skills we need are
story telling skills
not data mining
skills*

Clive Humby
DE/2014 Keynote



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

THE NARRATIVE

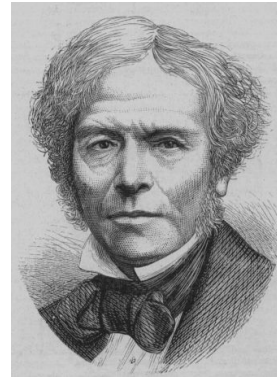
05/12/2014

Beyond Open Access

ON BEING A SCIENTIST

A GUIDE TO RESPONSIBLE CONDUCT IN RESEARCH

T H I R D E D I T I O N

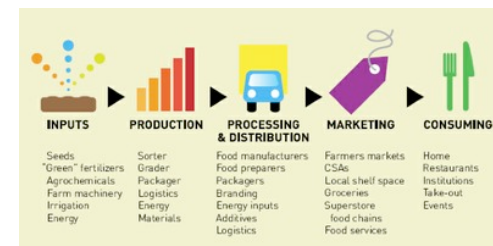


Faraday's laboratory notebooks are also remarkable in the amount of detail that they give about the design and setting up of experiments, interspersed with comments about their outcome and thoughts of a more philosophical kind. All are couched in plain language, with many vivid phrases of delightful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



WWW

THE INTERNET AND THE WEB

05/12/2014

Beyond Open Access

<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 boring!

Nature Chemistry

Web 2.0 & Semantic Web

USER GENERATED CONTENT

QUALITY?

05/12/2014

Beyond Open Access



Accuracy

<http://www.opsrules.com/supply-chain-optimization-blog/bid/282916/Why-You-Shouldn-t-Waste-Effort-Improving-Forecast-Accuracy>

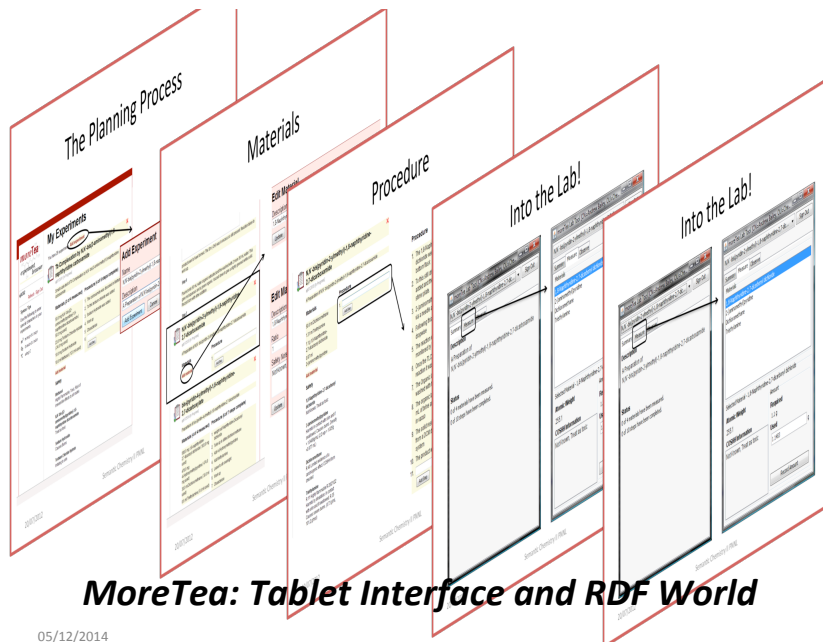
05/12/2014

Beyond Open Access

DIGITAL RESEARCH NOTEBOOKS

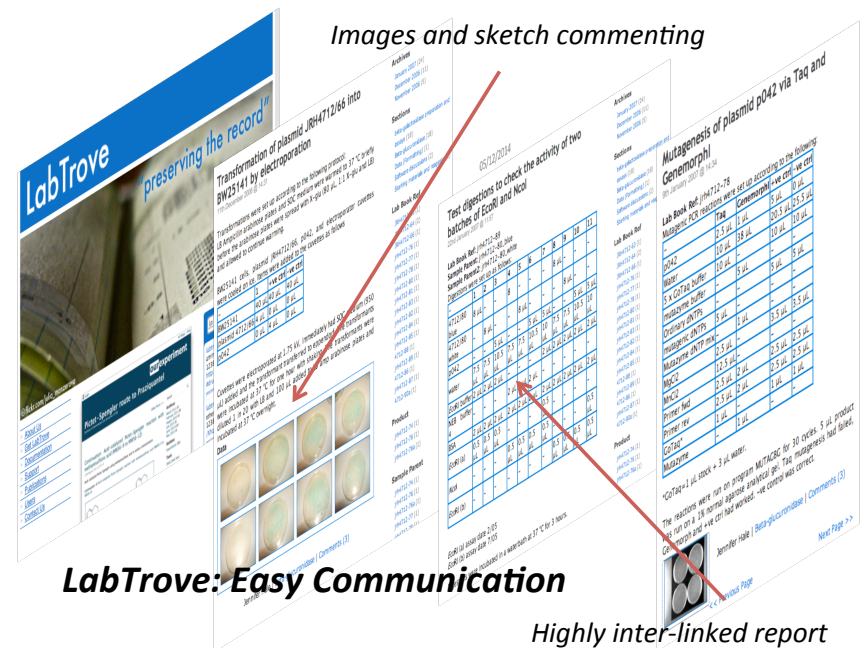
05/12/2014

Beyond Open Access



MoreTea: Tablet Interface and RDF World

05/12/2014



LabTrove: Easy Communication

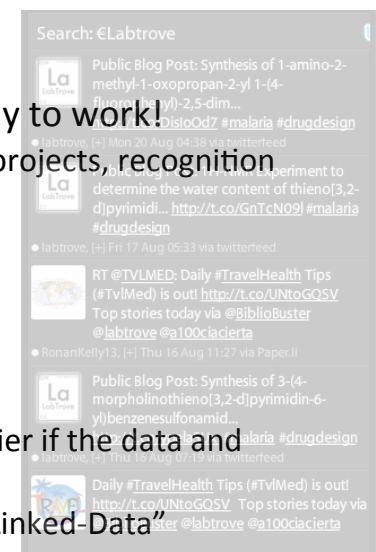


Digital and Physical

Open Notebook Science

- Certainly not always the way to work!
 - IPR, Commercial, long term projects, recognition issues, etc
- But
 - Makes connection much easier if the data and processes are "Open"
 - Easy to export & access of "Linked-Data"

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





LabTrove Open Notebooks Mat Todd's Malaria Project

The screenshot displays the LabTrove Open Notebooks interface. It features several tabs and sections: 'Pictet-Spengler route to Praziquantel', 'Racemization of PZQ and PZQamine', 'Hydrolysis of PZQ', and 'Monitoring the Reaction Progress by TLC'. Each section includes detailed chemical structures, reaction conditions, and progress monitoring data. The interface is designed to be user-friendly and accessible, with clear headings and organized content.

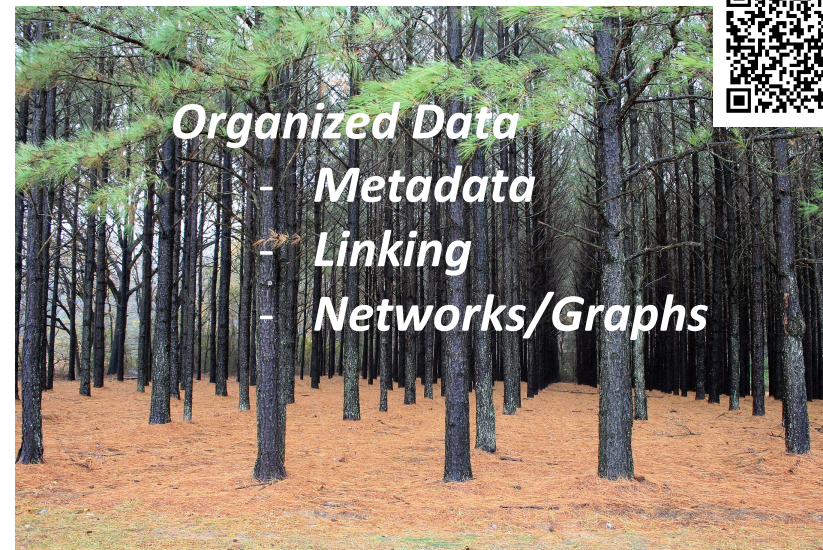
LabTrove -> RSS -> Email / Twitter

The screenshot shows the 'My Recipes' page on the LabTrove website. It features a navigation bar with 'My Recipes', 'Browse', 'Channels', and 'jgfray'. Below the navigation bar, there are tabs for 'Personal', 'Shared', and 'Favorites'. The main content area displays two RSS feeds: 'if [RSS icon] then [Email icon]' and 'if [RSS icon] then [Twitter icon]'. Each feed includes a description of the feed and a 'Create a Recipe' button. The page is designed to be user-friendly and accessible, with clear headings and organized content.



Finding the "Wood"
From the "Trees"

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



Organized Data
- Metadata
- Linking
- Networks/Graphs

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

05/12/2014

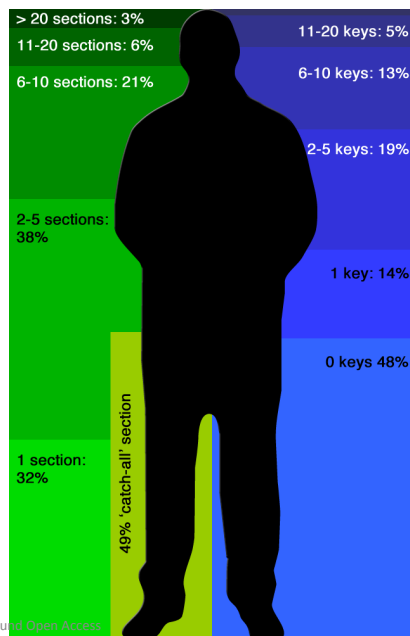
Beyond Open Access



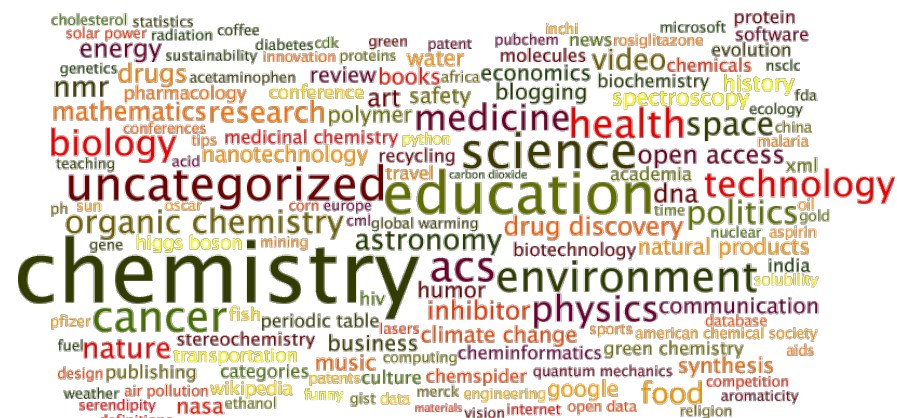
A big proportion of our users are not really adding metadata!

05/12/2014

Beyond Open Access



Chemistry blogs results



05/12/2014

Beyond Open Access

Categorizing 'activity' nouns



05/12/2014

Beyond Open Access

NASA results



05/12/2014

Beyond Open Access

Metadata

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

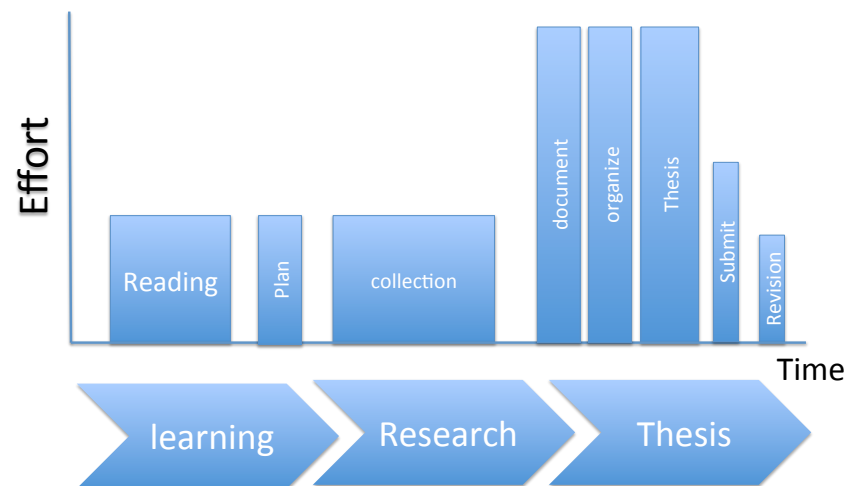
Toki Pona



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

05/12/2014

Beyond Open Access



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

05/12/2014

Beyond Open Access

Smart Tea Project

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 \(pdf download\)](#)
– The Royal Society of Chemistry provide the "definitive recipe for a cup of tea"

[British Standards Institute](#), 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)

P2 - Plans & Provenance

05/12/2014

Beyond Open Access

Graph/Network provides intuitive navigation

Depositor: [Simon J. Coles](#)
Structure: [Simon J. Coles](#)
Data collection: [Simon J. Coles](#) and [Michael B. Hursthouse](#)

Stages (10)

Stage	File
1	plan.pdf
2	plan.pdf



05/12/2014

Beyond Open Access

“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 re-use of data... the digital research notebook

DISSEMINATION IS PART OF THE RESEARCH

05/12/2014

Beyond Open Access

“Engrooved*”

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



<http://www.todayandtomorrow.net/2010/02/22/quarter-mile-groove/>

05/12/2014

Beyond Open Access

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**

05/12/2014

Beyond Open Access



Re-imagining Scientific Research

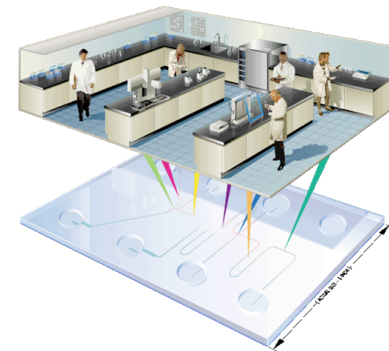
A “Social Machine” of experienced
researchers, technology &
information

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*

<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



Public bi

The samples you need may already be available

05/12/2014

Beyond Open Access

IDCC 2014 Meeting San Francisco

<http://www.dcc.ac.uk/events/idcc14/programme-presentations>

Got Data? Building a Global Ecosystem for Data-Driven Innovation
Dr. Francine Berman
Chair, Research Data Alliance / US
Edward P. Hamilton Distinguished Professor in Computer Science, RPI

Big Data in Biomedicine: Discovering new drugs and diagnostics from a trillion points of data
Lucile Packard Children's Hospital AT STANFORD
Atul Butte, MD, PhD
Chief, Division of Systems Medicine, Department of Pediatrics, Genetics, and by courtesy, Medicine, pathology, and Computer Science
Center for Pediatric Bioinformatics, LPCH
Stanford University
@atulbutte

Butte Lab

05/12/2014

Beyond Open Access

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/1.164291/menu/main/topColumns/topLeftColumn/pdf/516131a.pdf

05/12/2014

Beyond 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 untying 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*.

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

He who asks a question is a fool for five minutes;

he who does not ask a question remains a fool forever.

—Chinese Proverb

Quora

Search for questions, people, and topics

SHARE QUESTION

Follow Question

RELATED TOPICS

Macroeconomics

Monetary System

Currencies

Virtual Currency

Credits (Quora feature)

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

Follow Question 54

ASK TO ANSWER

ANSWERS

Quora User, Quora Admin

20 upvotes by Mike Probst, Quora User, Quora User, (more)

The reasons are quite straightforward:

The average return from:

- A question = 20 * Number of followers
- An answer = 10 for OP's upvote + 10 * Number of upvotes

Now (and this is a *post hoc* assumption) the average number of upvotes per answer is more than the average number of followers per question, so it makes sense to incentivize question addition by giving more weightage to number of followers.

You should also remember that questions drive Quora; without interesting questions there'd be no interesting answers.

Updated 30 Dec, 2013.

Upvote 45 Downvote Comments 36

Related Questions

Will Quora's credits result in more mundane questions being answered?

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?

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>

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