Open Access, Science Commons

and Open Science

Quoting from the Science Commons website (http://sciencecommons.org) :

“The time it takes to go from identifying a gene to developing a drug currently stands at 17 years — forever, for people suffering from disease."

"There are terabytes of research data being produced in laboratories around the world, but the best web search tools available can’t help us make sense of it.”

“Why? Because more stands between basic research and meaningful discovery than the problem of search.”

- **Copyrights for published data (Open Access)**
- **Open distributed collaboration (Open Science)**
- **Knowledge fragmentation and data standards (Science Commons)**

Copyrights for published data

As part of the publication process you surrender copyright to the journal over all the material of your manuscript (by signing a "Copyright Transfer Agreement").

An article by Peter Murray-Rust in Nature Precedings tittled "Open Science" (January 2008 - http://precedings.nature.com/documents/1526/version/1)

Case of a student that posted on her blog (= web page) a graph from a manuscript

What followed her post was notices for legal action from the journal, based on the claim of copyright violations !

Fundamental of science: open with research results, share, get/provide feedback

Journal copyright: do not to serve the interests of the scientist, but protect the profits of the publishing house.

Broadcasting of the article content on any medium is not allowed.

Then we all perform infringement each time we use in our presentations a graph from a published paper that is copyrighted !!!

Suppose you see a pattern in the data which the authors haven't observed (or maybe they noticed but did not want to write about ?)

How will your observation reach people's ears-eyes :

-- Publish a new article about it -> for a single observation ???

-- Write a letter to the journal-> maybe they publish, not much space to discuss

-- Put it on the web -> reach many people, open access, space for lots of discussion

But watch out for the lawyers !

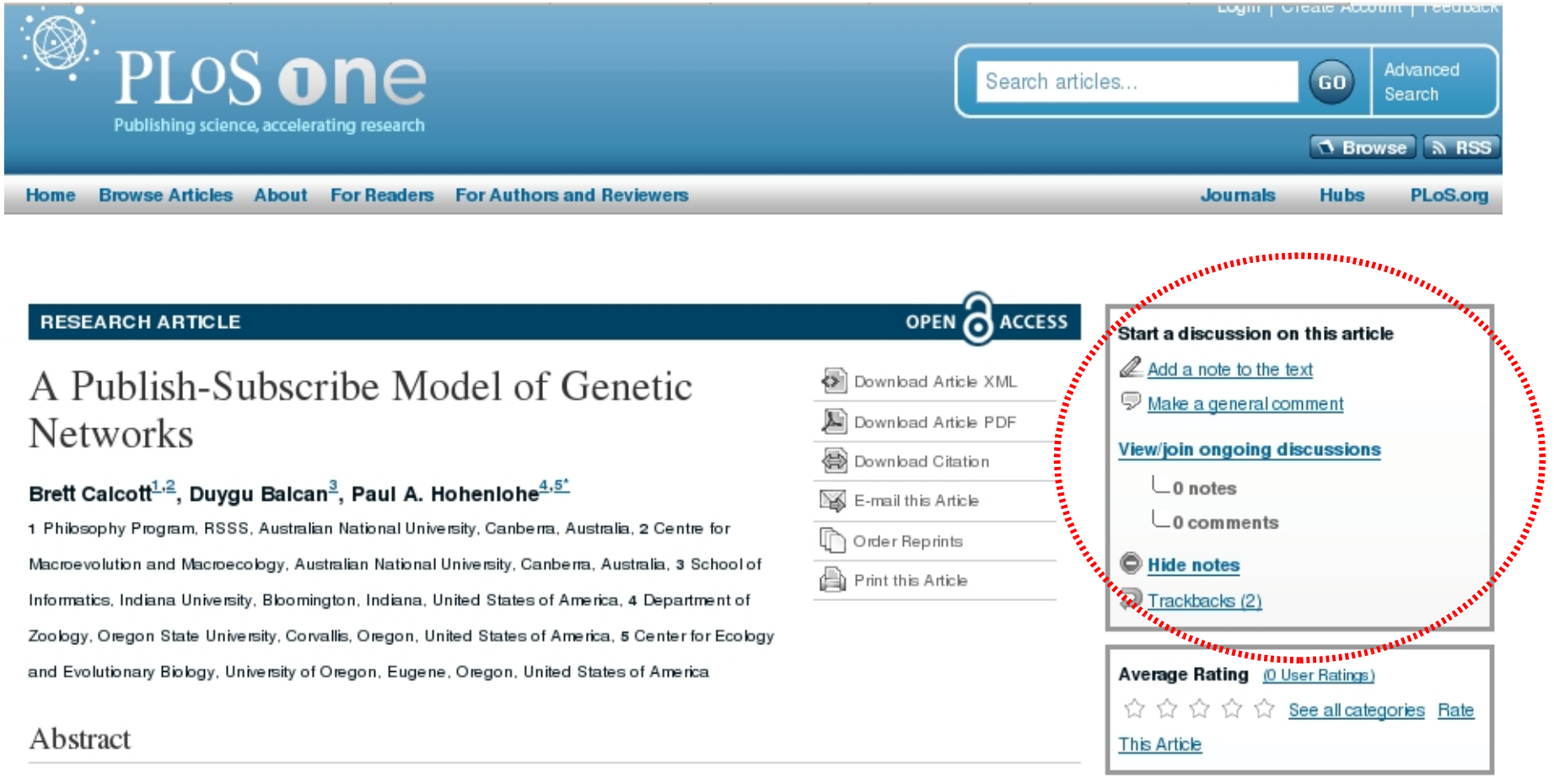
In the current Internet era, copyrights stifle progress

Audience free to link, read, post and discuss, data-mine the content of an article

Benefits the authors, feedback / discussion / peer-review on large scale

Open access = equal opportunity (pricey library subscriptions/developed nations)

Web 2.0 offers the medium and tools for this :



More Examples

nature precedings
Pre-publication research and preliminary findings

Log in
Register

Home | Latest Documents | Search: All documents | GO

- Bioinformatics (184)
- Biotechnology (103)
- Cancer (48)
- Chemistry (82)
- Developmental Biology (24)
- Earth and Environment (77)
- Evolutionary Biology and Ecology (153)
- Genetics (79)
- Immunology (39)
- Microbiology (49)
- Molecular Cell Biology (116
- Neuroscience (131)
- Pharmacology (46)

Contribute, comment and collaborate

Submit a document

Active Discussions

Embedding mac
readable prote
interactions da
scientific article
access and retr
Paolo Tieri et al.

1 comment, last about

Latest documents

arXiv.org > cs > cs.DS

Data Structures and Algorithms

Authors and titles for recent submissions

- Thu, 9 Oct 2008
- Tue, 7 Oct 2008
- Mon, 6 Oct 2008
- Thu, 2 Oct 2008
- Tue, 30 Sep 2008

[total of 10 entries: 1-10]
[showing up to 25 entries per page: fewer | more]

Thu, 9 Oct 2008

[1] arXiv:0810.1355 [ps, pdf, other]

Community Structure in Large Networks: Natural Cluster Sizes
Jure Leskovec, Kevin J. Lang, Anirban Dasgupta, Michael W. Mahoney
Comments: 66 pages, a much expanded version of our WWW 2008 paper
Subjects: **Data Structures and Algorithms (cs.DS)**; Physics and Society (physics.soc-ph)

Tue, 7 Oct 2008

[2] arXiv:0810.0906 [ps, pdf, other]

A linear time algorithm for L(2,1)-labeling of trees
Toru Hasunuma, Toshimasa Ishii, Hirotaka Ono, Yushi Uno
Comments: 18 pages (Title 1 page, Main part 10 pages, References 1 page, Appendix 6 pag
Subjects: **Data Structures and Algorithms (cs.DS)**

Cornell University Library

arXiv.org

Open access to 501,094 e-prints in Physics, Mathematics, Computer Science, Qua
Subject search and browse: Physics | Search | Form Interface

3 Oct 2008: arXiv passes half-million article milestone
See cumulative "What's New" pages.
Robots Beware: indiscriminate automated downloads from this site are *not* permitt

Physics

- Astrophysics (astro-ph new, recent, find)
- Condensed Matter (cond-mat new, recent, find)
  includes: Disordered Systems and Neural Networks; Materials Science; Meso
- General Relativity and Quantum Cosmology (gr-qc new, recent, find)
- High Energy Physics - Experiment (hep-ex new, recent, find)
- High Energy Physics - Lattice (hep-lat new, recent, find)
- High Energy Physics - Phenomenology (hep-ph new, recent, find)
- High Energy Physics - Theory (hep-th new, recent, find)
- Mathematical Physics (math-ph new, recent, find)
- Nuclear Experiment (nucl-ex new, recent, find)
- Nuclear Theory (nucl-th new, recent, find)
- Physics (physics new, recent, find)
  includes (see detailed description): Accelerator Physics; Atmospheric and Oc
  Statistics and Probability; Fluid Dynamics; General Physics; Geophysics; His
- Quantum Physics (quant-ph new, recent, find)

Mathematics

- Mathematics (math new, recent, find)
  includes (see detailed description): Algebraic Geometry; Algebraic Topology;
  Systems; Functional Analysis; General Mathematics; General Topology; Geo
  Theory; Numerical Analysis; Operator Algebras; Optimization and Control; Pr

Nonlinear Sciences

- Nonlinear Sciences (nlin new, recent, find)
  includes (see detailed description): Adaptation and Self-Organizing Systems;

Computer Science

- Computing Research Repository (CoRR new, recent, find)
  includes (see detailed description): Architecture; Artificial Intelligence; Compu
  Computer Vision and Pattern Recognition; Computers and Society; Cryptogra
  Literature; Graphics; Human-Computer Interaction; Information Retrieval; Info

Open distributed collaboration (Open Science)

group of minds can reach faster a solution, harness collective intelligence

prime example: Linux and all open source software

Linus invented besides the OS, a new system of mass collaboration

Open Innovation, search of skills further than local talent pool (http://en.wikipedia.org/wiki/Open_innovation)

Web 2.0 difference of a wiki or blog from a traditional web-page :

--> everyone can edit through browser, no need to maintain HTML/server

academia promotion to tenure: number of peer-reviewed publications

no metric system that provides accreditation to online distributed collaboration

page | talk | view source | history

OpenWetWare
Share your science.

navigation
- Main Page
- Recent changes
- Help

research
- Materials
- Protocols
- Resources

search
Go | Search

toolbox
- What links here
- Related changes

OPEN WETWARE

OpenWetWare is an effort to promote the sharing of information, know-how, and wisdom among researchers and groups who are working in biology & biological engineering. **Learn more about us.**
If you would like edit access, would be interested in helping out, or want your lab website hosted on OpenWetWare, please join us.

Labs & Groups From around the world | **Courses** Host & view classes | **Protocols** Share techniques & more | **Blogs** Read OWW blogs

Welcome
2008 IGEM Teams

OpenWetWare
Lab Notebooks
New features include:
Dynamic calendars
Create or view entries with a click

OWW Community Blog

Springer buys open access publisher BioMed Central

According to Scientific American, the not-so-open-access publisher Springer has just acquired the very open access pioneer site BioMed Central.

Charts and plots on OWW made easy

history

OPEN WETWARE

Courses Host & view classes | **Protocols** Share techniques & more | **Blogs** Read OWW blogs

In vivo

Arabidopsis thaliana [hide]
- Arabidopsis gDNA isolation
- Transformation Protocol for Arabidopsis

Escherichia coli [show]

Bacillus subtilis [show]

Mesoplasma florum [hide]
- Culture media
- Mesoplasma florum:Microarray design
- Mesoplasma florum:Codon usage
- Mesoplasma florum:Physiology
- Mesoplasma florum:Genome

Mouse [show]

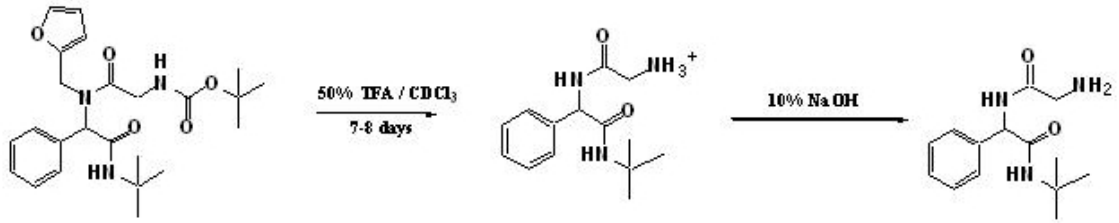
Mycobacterium smegmatis [show]

In vitro

Nucleic acids [show]

DNA [hide]
- 'Round-the-horn site-directed mutagenesis
- Affymetrix DNA labelling for gene expression arrays
- Agarose gel electrophoresis
- Anneal Extend PCR Protocol - for linking 2 pieces of DNA
- Annealing and primer extension
- Annealing primers
- Assembly pcr
- Bioanalyzer
- Centrifugal filtration/Nucleic acids
- DNA Precipitation
- DNA Synthesis from Oligos
- DNA ligation
- Designing primers
- Electrophoretic mobility shift assay

Exp101 | Protected | page | discussion | history | notify me



Objective

To isolate the product after a prolonged treatment of 99C with 50% trifluoroacetic acid in CDCl3 in order to confirm the Missing Methyl Mystery Mechanism also discussed as the Disappearing Methyls puzzle in the UsefulChem blog

Procedure

To a solution of ugi product 99C (88.6 mg, 0.2 mmol) in CDCl3 (500uL), trifluoroacetic acid (500uL) was added in an NMR tube. The reaction was monitored by NMR. After a complete conversion the solution was poured into a centrifuge tube and neutralized with 10% NaOH solution, extracted from with dichloromethane, dried with anhy. MgSO4. The product was subjected to NMR analysis after .

Results

Ugi Product (99C from Exp099)
H NMR
101A
t = 06min

Jean-Claude Bradley: open notebook in chemistry

Bora Zivkovic: first to blog unpublished scientific hypotheses and data

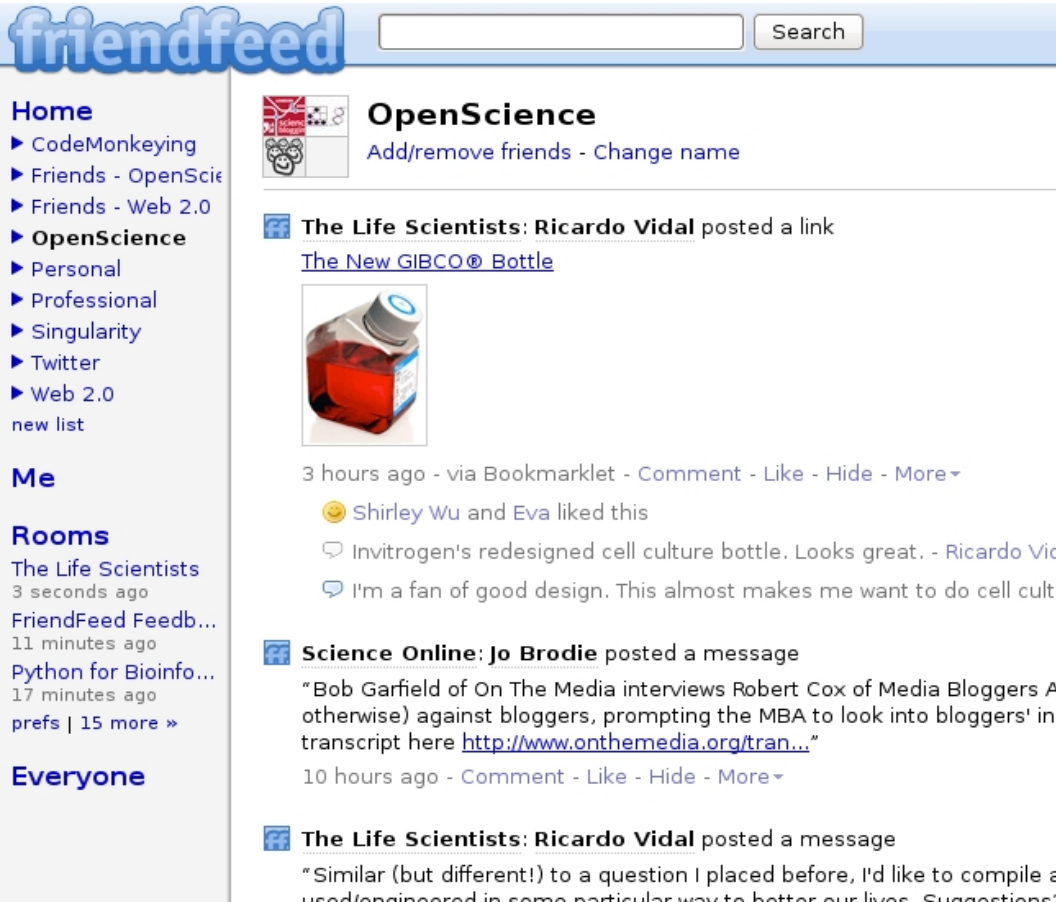
Rosie Redfield: blogging her research

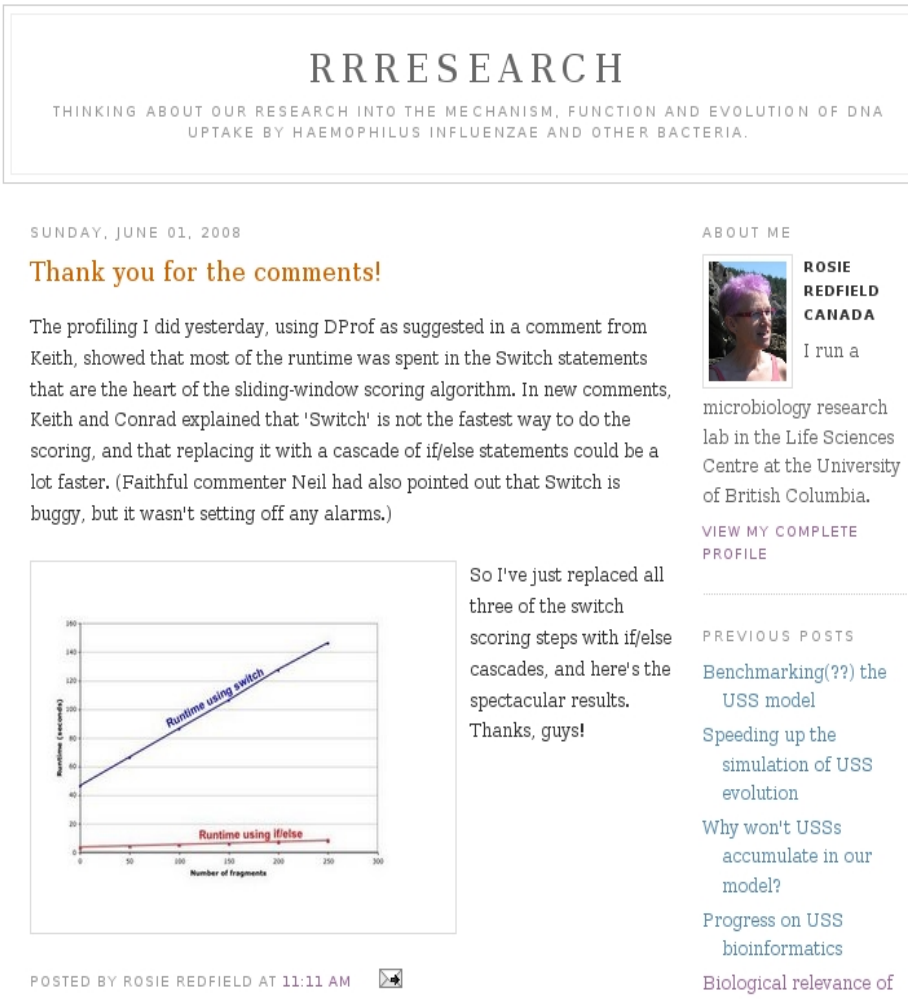
Reed Cartwright: being offered authorship after blogging about one paper

C.Neylon and P. Szczesney: molecular model for a grant through FriendFeed

John Wilkins: publishing a paper that started as series of blog posts

Jean-Claude Bradley: chemistry wiki





Knowledge fragmentation and data standards (Science Commons)

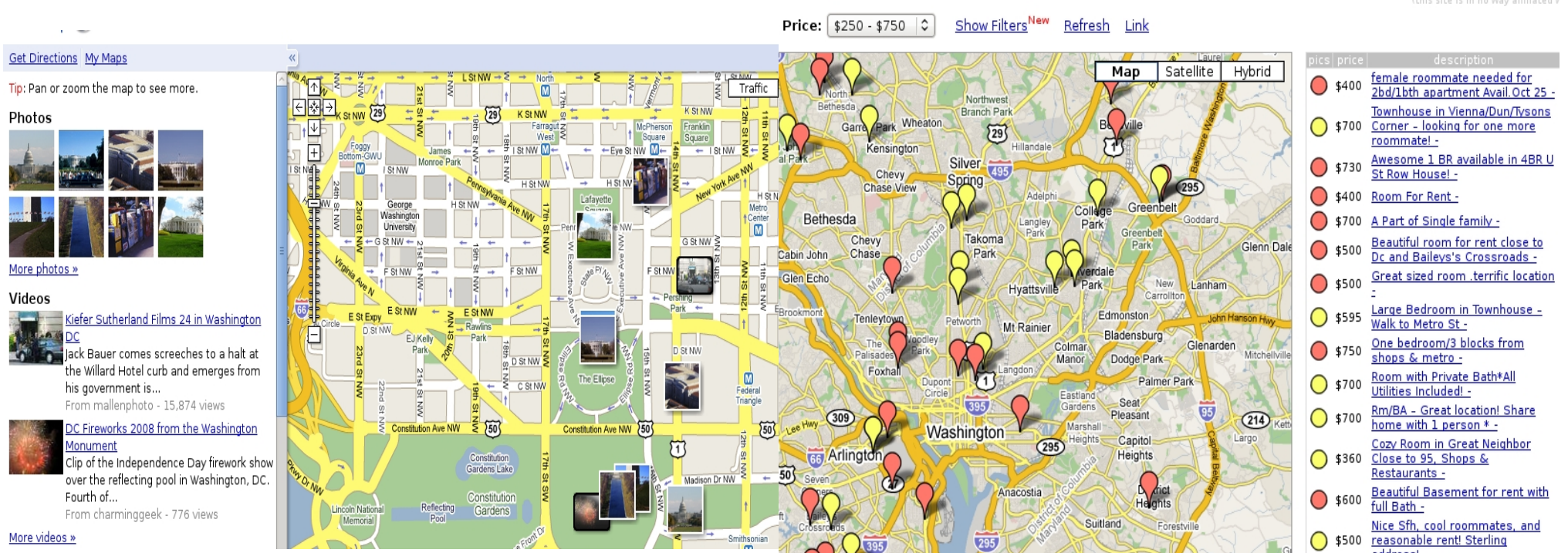
data locally deposited in the computers of research groups

or “available upon request” -- *on-the-fly-just-designed-to-get-the-job-done* format

need **Open** and **re-usable** scientific data, standardized data formats

semantic web technologies, XML -- machine processable

added value through aggregation



The standards at work on the www !
(mashups)

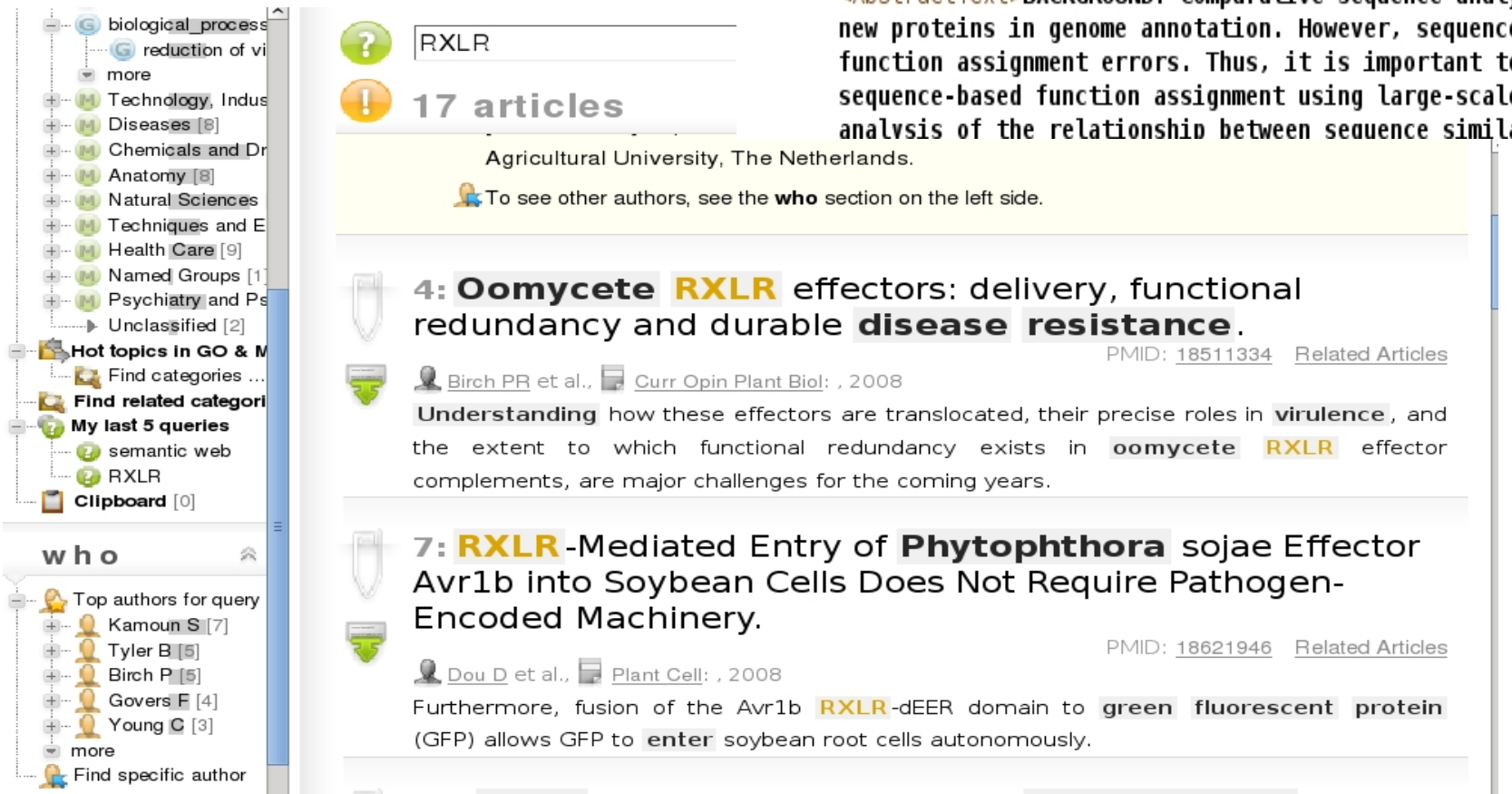
data standards – open access

easily integrating data from various sets of published experiments, new insights

also open access to standardized literature -- PubMed XML

imagine looking for a section that wassomewhere in those 50 papers I read last year.... :-(

```
<Article PubModel="Electronic">
   <Journal>
      <ISSN IssnType="Electronic">1471-2164</ISSN>
      <JournalIssue CitedMedium="Internet">
         <Volume>8</Volume>
         <PubDate>
            <Year>2007</Year>
         </PubDate>
      </JournalIssue>
      <Title>BMC genomics</Title>
      <ISOAbbreviation>BMC Genomics</ISOAbbreviation>
   </Journal>
   <ArticleTitle>Quantitative assessment of relationship between
   </ArticleTitle>
   <Pagination>
      <MedlinePgn>222</MedlinePgn>
   </Pagination>
   <Abstract>
      <AbstractText>BACKGROUND: Comparative sequence analysis is
      new proteins in genome annotation. However, sequence compar
      function assignment errors. Thus, it is important to perfor
      sequence-based function assignment using large-scale data i
      analysis of the relationship between sequence similarity ar
```



- **No Copyrights for published data (Open Access)**
- **Accreditation for Open distributed collaboration (Open Science)**
- **Data standards – value through aggregation (Science Commons)**

Thank you - time for Q & A

More detailed writings on these at my blog:

http://semanticlifescience.wordpress.com