Open Source Science

Maurizio Marchese
University of Trento, Italy

with the support of the LiquidPub project team
Challenge: doing science in the 21\textsuperscript{st}

- The Web has changed many fields:
  - News (blogs, RSS feeds, ...)
  - Music (p2p networks, iTunes, lastFM, ...)
  - Travel (Orbiz, Google maps,...)
  - Photos (Flikr, ...)
  - ...

- Has it changed also scientific knowledge production and dissemination processes?
Challenge: doing science in the 21\textsuperscript{st} century

• Yes! But - so far - mainly
  ▫ distributed working environment
  ▫ new and faster access channels
  ▫ …

• Scientific knowledge processes are still based on the traditional notion of “paper” publication and on peer review as quality assessment method
  ▫ *Philosophical Transactions of the Royal Society of London*—founded in 1665
  ▫ *Journal des scavans*—1665
  ▫ *Royal Society of Edinburgh’s Medical Essays and Observations*, - 1731, introduces peer review as we would recognise it today
We have a dream

Explore a real change of **paradigm**, culture and style of scientific production processes

Capture the lessons learned by the Web and open source, agile development to develop concepts, models, metrics, and tools for an efficient (for **people**), effective (for **science**), sustainable (for **publishers and the community**) way of creating, disseminating, evaluating, and consuming scientific knowledge.

From. www.52en.com/img/dream_01.jpg
Open Source Science  What is it?

• „Science“ („Research“)
  ▪ Advancing state-of-the-art of human knowledge
  ▪ Vs. Engineering (building a working solution)
  ▪ Vs. Innovation (bringing promising ideas to market)
Open Source Science. What is it?

• „Open“
  □ Science as a “Common”
    • scientific results belongs to everybody
    • when results are paid by the whole community, they have to be shared and kept in a common space
    • not limited to Open Access / Open Science
  □ Transparent processes
    • Common, agreed agenda – no hidden, private agenda
    • Open evaluation processes
Open Source Science. What is it?

• „Open Source“
  "I consider that the golden rule requires that if I like a program I must share its code with other people who like it and want to use it (Richard Stallman, 1983)"

  ▫ Not only „papers“
  • sharing artifacts (raw data, images, maps, videos, benchmarks, methods, toolkits, prototypes, infrastructures, ...)

  ▫ Collaborative, community-driven approach
  • e-Science initiatives
  • new way of creating, extracting and evaluating scientific work that mixes the expertise of the few with the “wisdom of the many”

  ▫ New ownership and credit attribution models
A large number of technologies are out there
<table>
<thead>
<tr>
<th>Blogs</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Science Blogs</td>
<td>Research Blogging</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wikis</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Google Docs Beta</td>
<td>Zoho Writer Beta</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Collaborative tagging and social bookmarking</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Connotea</td>
<td>delicious</td>
</tr>
<tr>
<td></td>
<td>social bookmarking</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scientific Search Services</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Google Scholar BETA</td>
<td>SCOPUS</td>
</tr>
<tr>
<td></td>
<td>citebase</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Journals with collaborative peer review processs</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PLoS one</td>
<td>nature proceedings</td>
</tr>
<tr>
<td></td>
<td>arXiv.org</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Complex Systems</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2collab</td>
<td>nature network</td>
</tr>
<tr>
<td></td>
<td>LiquidPub</td>
</tr>
</tbody>
</table>
But how they can be effectively used?

Let’s explore some dimensions of the issue

- Notion of “scientific contribution” and unit of dissemination
- Agile and Open Source “scientific” processes
- “Virtuous” metrics and efficient evaluation
Notion of “scientific contribution” and unit of dissemination

- In the past: papers
- What has been the function of the scientific paper?
  - The minimal scientific unit in which an idea or a result can be intelligibly expressed and evaluated
  - A rhetoric device that serves communication and pedagogical functions
Notion of “scientific contribution” and unit of dissemination

- However, many new things can be considered as contributions
  - Reviews, Comments, Blogs
  - Experiments, data
  - Prototypes, artifacts

- So the challenge is how to make them first class citizens in science (measure, search, disseminate....)
  - Define the new rhetorical models
  - How to reference them
  - How to evolve them
Agile and Open Source Development

- Scientific content generation and dissemination has mainly followed a “waterfall model”

Today we have new possibilities:

- Evolving, always in beta (liquid documents), with snapshots (solid documents)
- Incremental works are ok but should be recognized as such
- Errors and error corrections are also ok.
- Collaborative model: share early – in controlled fashion
- Continuous dissemination
Agile and Open Source Development

What do we need?

- Support definition and creation of “liquid” scientific contribution and their lifecycles
- Support tracing, support ownership, licensing, access right models
- Distributed and efficient search and navigation
- Challenges:
  - Usability, Light overhead
“virtuous” metrics and efficient evaluation

- metrics that take into account all different kinds of contributions.
  - bookmarking, forwarding,
  - Liquid Journals

- More fair and efficient peer review process

- Transparent metrics

“Not everything that can be counted counts, and not everything that counts can be counted.” -- Albert Einstein
Open / Transparent Review Processes

- Monitor and report
  - Fairness
  - Efficiency
  - Quality
How to make it happen

• **Understand culture first**
  ▫ Different disciplines have different (scientific, engineering, innovation) cultures
  ▫ Devise a discipline-specific approach

• **Have focused objectives**
  ▫ „Controlled experimentation“
  ▫ Mantain an holistic approach: technological, metodological, legal, economic

• **The Web matters**
  ▫ as technology
  ▫ as (social/scientific) collaboration platform
Thank you
Agile and Open Source Development

Many ICT tools available
- Version Control System
- CSCW Tools
- Web 2.0 Collaboration Tools

In progress
- Software prj mgm -> scientific artifact mgm
- Process workflow -> Scientific processes workflow