Force11: the Future of Research Communications and e-Scholarship

Anita de Waard
Disruptive Technologies Director, Elsevier Labs, Burlington, VT

Maryann E. Martone
University of California, San Diego

http://force11.org
Outline

• Science is becoming more distributed
• So we need to communicate better
• One initiative: Beyond the PDF
• Leading to Force11
• Plans & call for comments
Plants make data:

- Internet of things: we can interact with ‘objects that blog’ or ‘Blogjects’, that track where they are and where they’ve been;
- have histories of their encounters and experiences have agency
- have a voice on the social web
Larry Smarr makes lots of data:

- He wears:
  - A Fitbit to count his every step
  - A Zeo to track his sleep patterns
  - A Polar WearLink that lets him regulate his maximum heart rate during exercise
  - 23andMe analyzed his DNA for disease susceptibility.

- Your Future Health analyzed blood and stool samples for 100 biomarkers:
  - At one point, C-reactive protein stood out as higher than normal.
  - A blood test showed that his CRP had climbed to 14.5 during the attack.
  - He took antibiotics, the symptoms resolved, and his CRP dropped to 4.9—but that was still unusually high.
  - Lactoferrin, too, rose several times to sky-high levels—200, whereas the normal count is less than 7.3—and in tandem with CRP
  - Smarr now thinks his diverticulitis attack was actually Crohn's disease—and his gastroenterologist (reluctantly) agreed.
Clearity Foundation:
A translational medicine and public service foundation for:

- Providing doctors access to molecular profiling for their ovarian cancer patients
- Providing doctors and patients clinical trial options informed by individual tumor biology
- Providing financial support for the profiling work for patients – Oprah approved!
Brittany Wenger uses it:

Winner of the Google Science Fair 2012

17-year old Brittany Wenger developed a **cloud-based neural network** that is able to **seamlessly and accurately assess tissue samples** for signs/evidence of breast cancer to give more credence to the currently used (less reliable) minimally invasive procedure called Fine Needle Aspirates (FNAs).

By looking at nine different input features and comparing them to the training examples, Brittany’s cloud-based neural network **can detect malignant breast tumors with an accuracy of 99.11%**.

Because her neural network is deployed **in the cloud** using Google’s app engine it means it **can be accessed from existing medical systems as well as through a web browser or mobile apps.**
Mark Wilkinson uses it:

Given a protein P in Species X:

Find proteins similar to P in Species Y
Retrieve interactors in Species Y
Sequence-compare Y-interactors with Species X genome

(1) → Keep only those with homologue in
Find proteins similar to P in Species Z
Retrieve interactors in Species Z
Sequence-compare Z-interactors with (1)

→ Putative interactors in Species X

Using what is known about interactions in fly & yeast, predict new interactions with a human protein –

Running over data on the web that he neither created nor knew about!
Science is becoming distributed:

Data

Tools

Thoughts
Science is becoming distributed:

Data is king!
- Data needs to say what it’s about
- Data needs to say where it comes from
- Data needs to know who owns it
- Data needs to be sensitive to privacy
- Data needs to know how it’s used
Science is becoming distributed:

Tools rule!
Tools can be made by everyone:
Tools are open and free
Tools will know where data lives
Tools need to know about data:
• Privacy/ownership
• Trustworthiness
• Provenance
Science is becoming distributed:

If data and tools are ubiquitous, what matters most are the questions you ask:
• What is interesting?
• What is important?
• Who cares?
1. Add metadata to everything
2. Use a workflow tool
3. Write in a shared space
4. Invite reviews
5. The reviewer approves (or comments, author revises, etc)
6. Run nifty apps over all of this.

Rats were subjected to two grueling tests (click on fig 2 to see underlying data). These results suggest that the neurological pain pro-

Calculate, coordinate...
Compile, comment, compare...

Review  Edit  Revise
What do we need to get there?

• 1. **Metadata standards**: Standards that allow interoperable exchange of information on any knowledge item created in a lab, including provenance and privacy/IPR rights

• 2. **Tools**: *Workflow tools* that work for all science, are scalable, safe, and user-friendly

• 3, 4, 5. Semantic/Linked Data-Centric **authoring, annotation and editing** environments that enable interlinked, distributed knowledge creation.

• 6. Publishing systems that run as **application servers**.

=> **Social change**:

– Scientists need to realize they should annotate their work
– Libraries change their visions and jobs
– Publishers realize they need to take on new roles
Beyond the PDF
Jan 2011 San Diego

Application of emergent technologies to measurably improve the way that scholarship is conveyed and comprehended
Outcome of Beyond the PDF:

• Community interested in connecting
• Topics:
  – New formats for the research paper
  – Tools for creating, (re)viewing, assessing, editing
  – Connecting workflows and data to papers
  – New metrics for success
  – New business models?
• Some discussion; many initiatives- no real coordination
• Forc: how do we take this a step further?
Future of Research Communications:

Many workshops, papers, conferences, meetings, reports, about innovation in science publishing:

Many great ideas, but still a lack of large-scale change
Some arguments: ‘I can’t get funded for that’, or ‘the publishers will never agree to that’ or ‘the reward system is just not set up that way’ or ‘my university/dean/provost doesn’t believe in it’
Here (hopefully) the people you are pointing at are in the room!
FoRCe11 at Dagstuhl
The Manifesto

Bourne, Philip E.; Clark, Timothy W.; Dale, Robert; de Waard, Anita; Herman, Ivan; Hovy, Eduard H.; Shotton, David
Weitere Beteiligte (Hrsg. etc.): Philip E. Bourne and Timothy W. Clark and Robert Dale and Anita de Waard and Ivan Herman and Eduard H. Hovy and David Shotton

Improving The Future of Research Communications and e-Scholarship (Dagstuhl Perspectives Workshop 11331)

Abstract

The dissemination of knowledge derived from research and scholarship has a fundamental impact on the ways in which society develops and progresses, and at the same time feeds back to improve subsequent research and scholarship. Here, as in so many other areas of human activity, the Internet is changing the way things work; two decades of emergent and increasingly pervasive information technology have demonstrated the potential for far more effective scholarly communication. But the use of this technology remains limited. Force11 is a community of scholars, librarians, archivists, publishers and research funders that has arisen organically to help facilitate the change toward improved knowledge creation and sharing. This document highlights the findings of the Force11 workshop on the Future of Research Communication held at Schloss Dag...
Core issues of Force11 Manifesto

<table>
<thead>
<tr>
<th>Problems</th>
<th>Recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Formats and Technologies</strong></td>
<td></td>
</tr>
<tr>
<td>2.1 Existing formats needlessly limit, inhibit and undermine effective knowledge transfer</td>
<td>3.1 Rethink the unit and form of the scholarly publication</td>
</tr>
<tr>
<td>2.2 Improved knowledge dissemination mechanisms produce information overload</td>
<td>3.2 Develop tools and technologies that better support the scholarly lifecycle</td>
</tr>
<tr>
<td>2.3 Claims are hard to verify and results are hard to reuse</td>
<td>3.3 Add data, software, and workflows into the publication as first-class research objects</td>
</tr>
<tr>
<td><strong>Business Models and Attribution of Credit</strong></td>
<td></td>
</tr>
<tr>
<td>2.4 There is a tension between commercial publishing and the provision of unfettered access to scholarly information</td>
<td>3.4 Derive new financially sustainable models of open access</td>
</tr>
<tr>
<td>2.5 Traditional business models of publishing are being threatened</td>
<td>3.5 Derive new business models for science publishers and libraries</td>
</tr>
<tr>
<td>2.6 Current academic assessment models don’t adequately measure the merit of scholars and their work over the full breadth of their research outputs</td>
<td>3.6 Derive new methods and metrics for evaluating quality and impact that extend beyond traditional print outputs to embrace the new technologies</td>
</tr>
</tbody>
</table>
FORCE11: Mission

Force11 is a community of scholars, librarians, archivists, publishers and research funders that has arisen organically to help facilitate the change toward improved knowledge creation and sharing. The manifesto ... summarizes a number of key problems facing scholarly publishing today, and presents a vision that addresses these problems, proposing concrete steps that key stakeholders can take to improve the state of scholarly publishing.

http://force11.org
Phil Bourne requested and obtained funding for 2012 from the Sloan Foundation to take this to the next step.

Goals:
- Improve collaborative practice among stakeholders
- Coordinate standard and technology development
- Advocate for advancing scholarship across multiple venues
- Co-develop proposals to effect change

Progress:
- Website established; make use of social media (Twitter, Google)
  - >225 members
  - Active outreach
- Resource repository (tools, data, projects, materials) and blogs, relevant papers
- Planning for the next Beyond the PDF conference
- Outreach to increase awareness and broaden representation
We will invent the future...

- Like Larry’s quantified self, scientists have ways of exposing their expertise and products on the web unfiltered
  - Blogs, videos, data sets
- The web leads to new metrics of impact
  - Connectivity, social presence
  - Altmetrics

A History of Scientific and Technical Periodicals: “The erudite letter was used as a form of exchange for ideas and news of the learned world as well as a form of primary ‘publication’. Leibniz, for example, wrote a complete treatise on philosophy in one series of letters….”4 (p. 56)
Beyond the PDF2

• Planning is underway for the next Beyond the PDF conference (March 19-20, 2013, Amsterdam)
  – New models of content creation
  – New models of content distribution
  – New models of evaluation
  – New business models for publishing
  – Making it happen
    • Challenges
    • Match-making
  – Beyond the horizon

• Join FORCE11 now (members get first chance to attend Beyond the PDF2)

http://www.force11.org/beyondthepdf2
FORCE11 as a catalyst for change

• What approaches to dissemination, review and assessment work?
  – What evidence do we have?
  – What should we adopt now?
• What tools, systems, and framework are needed to support new scholarship modes? Who pays for them?
• How do we persuade the research community to change aka “It’s a cultural issue...”

• How do we better involve and connect to the activities of CNI members?
• What do you need from FORCE11?
  – Users?
  – Tools?
  – Collaborators?
  – Advertising?
  – A bully pulpit?

http://force11.org