A quick trip through openness, freedom and transparency

An introduction for scientists ... and everybody else.

Konrad Förstner, Bork Group, EMBL

http://konrad.foerstner.org

October 11th, 2007
Disclaimer

- Fasten your seat belts! This will be a rushed journey through a lot of topics that have some important concepts in common.
- It is an introductory talk:
  - ⇒ just an appetiser
  - ⇒ some simplifications / idealisations / polarisations
- Sometimes techy
Starting consensus

- Openness
- Freedom
- Transparency

are essential for a functional and democratic scientific community and for society in general.
Open Source Software

What is it?

Open source software is computer software that is published under a license that gives the freedom to use, modify and redistribute it.
Open Source Software

Closed Source Software:
Only the executable program (binary) is available.

Open Source Software:
Source code and executable program are available.

```c
int main (void){
    int a = 23;
    int b = 42;

    printf("I have %d bananas.\n", a);
    printf("You have %d bananas.\n", b);
}
```
Open Source Software

FLOSS/FOSS/OSS licenses

- *(Free)/(Libre)/Open-Source Software*
- Make software “free as in freedom, not as in free beer”.
- Most important examples:
  - GNU General Public License (“copyleft”)
  - BSD License
Important freedoms that Open Source licenses give

- The freedom to use
- The freedom to copy
- The freedom to make derivatives
- The freedom to redistribute (also the modified versions)
Open Source Software

**Advantages**

- Transparency
- Potentially higher quality/security due to peer-review by the community
- Independence of vendors
- Modification/adaptation to personal needs possible
- Reusability of code ⇒ faster development
- Free/low costs ⇒ affordable for everybody
Open Source Software

Some examples you might be familiar with

- Mozilla Firefox (Web browser)
- OpenOffice (Office Suite)
- GNU/Linux (Operating system)
- Gimp (Graphics editor)

Some less obvious examples that you use for sure

- Apache (Web server)
- BIND (DNS server)
- Google (adapted version of GNU/Linux on their servers)
Open Source Software

Some examples you might be familiar with

- Mozilla Firefox (Web browser)
- OpenOffice (Office Suite)
- GNU/Linux (Operating system)
- Gimp (Graphics editor)

Some less obvious examples that you use for sure

- Apache (Web server)
- BIND (DNS server)
- Google (adapted version of GNU/Linux on their servers)
Open Formats/Standards

What is it?

Open formats are publicly available specification for storing digital data that are free of legal restrictions and free of charge.
Open Formats/Standards

An analogy – the ISO container
## Open Formats/Standards

### Examples

<table>
<thead>
<tr>
<th>Proprietary</th>
<th>Open</th>
</tr>
</thead>
<tbody>
<tr>
<td>doc, xls (also Office Open XML)</td>
<td>Open Document Format</td>
</tr>
<tr>
<td>ICQ, MSN, AIM</td>
<td>Jabber/XMPP (Google Talk)</td>
</tr>
<tr>
<td>Skype</td>
<td>SIP-based</td>
</tr>
<tr>
<td>luckily nothing</td>
<td>HTML/XHTML</td>
</tr>
</tbody>
</table>
Open Formats

Advantages

- User can choose the application to use – no vendor lock-in
- Possibility to write own applications using that format
- Due to above points safe to be usable also in future
Creative Commons licenses define the spectrum of possibilities between full copyright (*all rights reserved*) and the public domain (*no rights reserved*).
Creative Commons licenses

Why?
The default copyright (© all rights reserved) restricts creativity and cultural development in the digital age. Creative Commons licenses make it easy for creators to define the freedom of their creations.

Use for ...

- Text
- Images
- Audio
- Video
Creative Commons licenses

Select a license by choosing conditions

- Attribution
- No derivative works
- Non-commercial
- Share alike

Some examples of CC licenses

- People + Circle = Creative Commons Attribution-No Derivative Works
- People + Dollar = Creative Commons Attribution-Noncommercial
- People + Dollar + Circle = Creative Commons Attribution-Noncommercial-No Derivative Works
Open Access is a publishing concept with immediate, free and unrestricted online access to scholarly publications.
Open Access

Current closed-access scenario

1. Scientist is paid by the public to do research
2. Scientist writes publication and gives (nearly) all the rights to publisher
3. Other scientists (if they can) buy the publication with public money

⇒ public pays but loses rights to the publisher
⇒ access to knowledge is limited to people who can pay for it
Publishing with Open Access

- Publishing using an Open Access publisher
  - Author keeps rights (often a Creative Commons license is used)
  - Different business models: fee or non-fee-based
- Open access self-archiving
  - In addition to the traditional publication the article is archived in central repository (e.g. arXiv, Nature Precedings)

⇒ Immediate access to the generated knowledge for everybody
Open Access

Advantages

- Knowledge is not locked
- No financial hurdles for readers
- Higher scientific impact
- Computational text analysis possible (necessary due to growing amount of literature)
Sometime I have the feeling people forget ...

- that scientists are payed by the public to generate knowledge for the public.
- that delayed/restricted access to results/data/knowledge hampers scientific progress and maybe even costs lives.
- that “negative” results are also results.
- the question if the “paper” is really the optimal form of communicating science in the internet age.
Example 1: PloS One

- An online-only Open Access journal
- Pre-publication peer review but not filtered by scientific relevance (= don’t care about impact factor)
- Users can rate and discuss articles after publication
Example 2: Science Commons

- Aim is “removing unnecessary legal and technical barriers to scientific collaboration and innovation.”
- E.g. *The Biological Materials Transfer Agreement Project* (MTA): lower the costs of transferring physical biological materials (DNA, cell lines, model animals etc.)

- ... and family: Wikibooks, Wikiversity etc.
- An easy way of teaching a broad audience and communicate science to the public
Example 4: OpenWetWare

- “Sharing of information, know-how, and wisdom among researchers and groups who are working in biology & biological engineering.”
- Wiki-based platform
Example 5: Open Notebook Science

- Online version of the classical lab notebook
- Making the way of discoveries transparent
- Instant publication of results
- Instant feedback from colleagues
- Also negative results are presented
Example 6: Wikiscience

- Article hosted on a wiki
- All versions are stored
- Constantly improving
- Many contributors
- Micropublications
Yes, there are currently problems like

- Fear of being scooped as blog/wiki contributions are not official publications
- or not accepted by journals as already published in blogs/wikis
- No credit system for this kind of scientific contribution

⇒ Problems are cultural not technical!
⇒ Luckily there are grass root projects where people start to play around with the new concepts.
Take home messages

- More openness, freedom and transparency can improve a lot of fields.
- Science is one of them and you can help.
  - Learn
  - Test
  - Discuss
  - Spread the word
  - Question the current status
Acknowledgements

Thanks to …

- An uncountable amount of people who communicate, discuss and test the presented ideas
- Bernd Ahlers who ignited my Open Source fire
- Anne-Marie Glynn for being the test drive audience.
Let’s open it
Intro

Open Source
Open Formats
Creative Commons
Open Access
Open Science / Knowledge

THM

Selected references and Image sources/attribution

Reference

http://www.opensource.org
http://www.openformats.org
http://creativecommons.org/
http://www.plosone.org/
http://sciencecommons.org/
http://wikipedia.org/
http://openwetware.org
http://www.edge.org/3rd_culture/kelly06/kelly06_index.html

Image sources (The other images were created by myself)

Seatbelt sign http://flickr.com/photos/davescunningplan/236094576/ by Dave Spellman
Hanging fishes http://flickr.com/photos/rastafabi/369869352/ by rastafabi
Shell http://flickr.com/photos/96dpi/501424695/ by Andreas Levers
Freedom http://flickr.com/photos/dazzied/427180864/ by Dazzie D
Shipping containers http://flickr.com/photos/16543356@N00/150898441/ by melted_snowball
Container train http://flickr.com/photos/telstar/163503065/ by Todd Lappin
Container trucks on an American highway http://flickr.com/photos/87913776@N00/422603859/ by futureatlas.com

I have no idea why that caged bird does a damn thing http://flickr.com/photos/emdot/135529627/ by marya
Creative Commons Logo http://flickr.com/photos/purz1baum/239202519/ by Claudio Schwarz
Filters Showcase http://flickr.com/photos/denverjeffrey/300338550/ by Jeffrey Beall
Open Access Logo http://open-access.net/de/austausch/downloads/

Hardship in the streets of Varanasi (India) http://flickr.com/photos/ahron/266050467/ by Ahron de Leeuw
Tree of Knowledge http://flickr.com/photos/knilram/64366434/ by Knilram
Note Book http://flickr.com/photos/prashant_zi/289482096/ by Prashant ZI
tough grass http://flickr.com/photos/zachk/109921799/ by zach kowalczyk
Corkscrew http://flickr.com/photos/awrose/121085717/ by Adam Rose
A video of the talk can be found on http://konrad.foerstner.org

Created in \LaTeX\ using the \texttt{beamer} class, pdfl\LaTeX\ and \texttt{Emacs}.
\texttt{Gimp} and \texttt{Firefox} were used to take screen shots of websites.
All these programs run on \textit{OpenBSD}.

http://www.latex-project.org
http://latex-beamer.sourceforge.net
http://www.tug.org/applications/pdftex
http://www.gnu.org/software/emacs
http://www.gimp.org/
http://audacity.sourceforge.net/
http://recordmydesktop.iovar.org
http://www.openbsd.org

Published under the Creative Commons Attribution 3.0 License

http://creativecommons.org/licenses/by/3.0/

Document version 1.1 2007/10/11