Three Revolutions

George O. Strawn
NSF CIO

Note bene

This talk will outline the current National Science Foundation policies regarding scholarly publications that result from NSF-supported projects.

It will also seek to present an IT-based historical perspective on Web publishing. As such, it reflects the experiences and opinions of the presenter, not the policies of the National Science Foundation

Outline

- Some Books
- Four Renaissances
- Disruptive Technologies
- Three Revolutions
- Web Publishing

Some Books

- *The Printing Press as an Agent of Change*
  Eisenstein (1980)

- *Media, Technology and Society*
  Winston (1998)

- *The New Renaissance*
  Robertson (1998)

- *The Innovator's Dilemma*
  Christensen (1997)

Four Renaissances

(Robertson, The New Renaissance)

Years ago	*Technology*	*Enabled*
50,000	Spoken language	Culture
5,000	Written language	Civilization
500	Printed language	Science, nation state, Reformation, individual
50	Computed language	”Science++”, globalization

Sustaining and Disruptive Technologies

(Christensen, The Innovator's Dilemma)

- A *sustaining technology* provides an improved solution for a given problem

- A *disruptive technology* provides a poorer solution for a given problem; e.g.,

  – PCs for computing

  – The Internet for telecommunications

  – The Web for publishing

The Potential of a Disruptive Technology

- Since it starts out as a poorer solution, it may be ignored by incumbent providers and customers alike

- But if stays around as a solution to other problems (i.e., finds a niche)

- And if gets on a faster improvement track than the earlier technology

- It may then disrupt the incumbents (up to and including putting them out of business) when it becomes a better, cheaper solution

Revolution One: The PC

- Started literally as a "toy" for hobbyists and educators
- Was based on the rapidly improving integrated circuit technology (the computer chip)
- Incumbents (mainframe and mini makers) didn't like the low margins of PCs and did their best to ignore/suppress them
- All but one mainframe and mini-computer maker (IBM) were put out of business by PC makers who created business plans based on volume to survive the lower prices and margins

Revolution Two: The Internet

- The Internet began as a U.S. government-supported experiment in *packet switching* technology (as opposed to *circuit switching* technology, which the telecom industry utilized)

- *Fiber optic* communications, developed by the telecom industry to improve telephone service, together with computer chips, became the enabling technologies that made packet switching "faster, cheaper and better" than circuit switching

- Most telecom companies resisted the Internet, but the monopoly local phone service cash flow of the regional Bell operating companies (now consolidated into ATT and Verizon) helped them survive and then to buy or bury most Internet providers (e.g., PSInet, UUnet, MCInet)

Revolution Three: The Web

- The Web (WWW) was invented in 1990 as a means to share information over the Internet at CERN, the European Physics Lab
- The Mosaic browser was invented in 1993 at the NSF-supported NCSA at the University of Illinois
- Mosaic and subsequent browsers hid the complexity of the Web and made it the Internet information sharing method of choice (e.g., NSF initiated electronic proposal submission in 1995)
- Many “printing press products” are now challenged by Web alternatives

Web Publishing

- Initially a toy, like all disruptive technologies
- Now the Web provides a viable publishing model for
  - News (and newspapers are folding)
  - Encyclopedias (Wikipedia shouldn't work, but it does)
  - Books (from out-of-copyright freebies to *Kindle* purchases)
  - *Scholarly literature*
- A variety of business models are emerging
  - Google ads are e-versions of newspaper ads
  - Wikipedia is a volunteer labor of love
  - Many $25 hardbacks cost $9.99 on Kindle
  - Journal e-subscriptions, pay per e-article, author pays, funding agency pays, etc.

Open Access to Scholarly Publishing?

- NIH open access policy ("the public paid for it, the public should get to see it")
- Cornyn-Lieberman bill (open access for all Federal agencies)
- Conyers bill (open access for no Federal agencies)
- Harvard, MIT, Stanford and BU have announced open access initiatives (published faculty research to be placed on their *institutional repositories*)

Open Access--what's in it for:

- Authors: wider availability of their work
- Researchers: more access, esp. at small schools
- University Libraries: acquisition budget relief
- Universities: increased prominence
- Publishers: the need to create new business plans
- Scientific societies: also new business plans
- Science: acceleration in scientific progress
- The Public: more rapid technological progress

NSF Access Policies

- NSF has always “outsourced” publication
- NSF maintains a publicly accessible database of all funded projects (PI name, proposal abstract, etc)
- NSF has added publisher citations for the publications which have resulted from each project (5.1 publications per project, on average)
- NSF will add (part of) the new USG-standard final project report to this public database later this year (not a peer-reviewed report)

Open access is not free access

- Disruptive technologies produce cheaper solutions (which disrupt incumbents business plans)
  - PCs are much cheaper than mainframes
  - Internet costs are much cheaper than telephone costs (because of efficient use of communication channels)
  - Web publishing is much cheaper than printing press publishing
- So the question is not free access versus paid access: rather it is *cheap access versus expensive access*

Use of Copyright

- Publishers' copyright advantage is similar to the telecom companies' monopoly advantage (whereas mainframe and mini makers had no similar barrier to entry)

- Will incumbent publishers similarly use copyright to "hang on" until they can put in place practices and business models that embrace web publishing (i.e., drastically lower prices) and be survivors like the telecom companies?

Web publishing: What might it be like?

- Authors could submit articles directly to an "arXiv"

- There could be "approved" reviewers for each subject area, who may review and rate articles (public reviews and ratings could be accommodated separately)

- Reviews and ratings might be displayed like results of book reviews on amazon.com and could become the new *"million eyeballs" peer review system* (approved and non-approved reviews being displayed separately)

- The cost of such a system might be 100x less expensive than conventional peer reviewed journals

Web Publishing: New Services

- Publishers will facilitate annotating articles with metadata (e.g., Semantic Web rdf triples) to enable text mining

- Publishers will also facilitate generalizing the concept of an article to include the supporting data sets (which will be citable and also annotated to facilitate computer processing)

- Both text and data will be reuseable (with proper attribution) to facilitate creation of derived works

- These annotated articles and data will be "readable" by computers as well as humans. New science (especially interdisciplinary science) will be mined from such annotated articles

In Summary

- The PC revolution put incumbent computer makers out of business because they did not react fast enough to the revolutionary potential of the new chip technology

- The Internet revolution stressed incumbent telecommunication providers, but they had the resources to hang on until they could adapt and provide the revolutionary technology (along with new computer makers who claimed part of the new business)

- The Web publishing revolution is happening. Will publishers be casualties like mainframe makers or survivors like telecommunications companies?