Disruptive Technologies, Innovation, and Digital Libraries Research – The Case of a Billion-Dollar Business

Yi-Tzuu Chien
World Technology Evaluation Center, Inc
Baltimore, MD 21224, U.S.A.
www.wtec.org; ychien@wtec.org

Abstract

After nearly a decade of research in digital libraries (DL) since the first major government sponsored initiative (DLI) launched in the United States and other similar programs elsewhere, the public is beginning to see the many fruits of the collective accomplishments from this research. Innovations in technology, testbeds, and content-based tools rank among those that have resulted from the scientific progress made in that period. The road ahead, however, requires a broader and better connection of DL research to other technologies and to society. If DL is to become a viable field of study and a permanent institution, it has to be sustainable not only technologically but also socially, politically, and economically.

This paper makes the arguments of charting a more sustainable course for the future of digital libraries research. A key argument here is the need to create a content-based business model that makes effective linkages between the "disruptive" forces of enabling technologies and the changes of direction that must be made in future DL research. How to make those linkages are the main challenges of the DL community for the coming decade.

Keywords: Digital library, roadblocks, innovation, business model, disruptive technology

1. Introduction

Around the time the Internet and the World Wide Web became popular in the early 1990s, it also became clear that the role of "information" or "electronic content" will be increasingly important in making the computational and network infrastructures more useful to the society. The High Performance Computing and Communications (HPCC) program in the United States recognized the significance of this shift and established a task force on "Information infrastructure technology and applications" (IITA). A goal of the IITA task force is to develop a research agenda and implementation plan for an expanded HPCC by adding a new "information" component. After months of deliberation, this task group published a report in February 1994, entitled appropriately, "Information Infrastructure Technology and Applications"[1]. This report became an influential document that helped re-define the HPCC program and subsequently its successor program. In that report, "Digital Library" was listed as one of the nine major "national challenges", defined as "information-intensive applications that have broad and direct impact on the nation's well-being and competitiveness". The report also outlined a general vision and research agenda for Digital Libraries. It stated (p. 18 of [1]):

"Digital Library includes work in both technologies and applications which will lead to significant advances in the generation, storage, and use of digital information of different kinds across high speed networks. A digital library is a knowledge center without walls, open 24 hours a day and accessible by way of a network. Research areas range from advanced mass storage, online capture of multimedia date, intelligent filtering, knowledge navigation, effective user interfaces, system integration, to prototyping and technology demonstration".

These words indeed sowed the early seeds for a major Federal R&D effort known as the Digital Library (DI) research initiative in the United States. Many other programs at the national and local levels have been launched since that time around the world. In the ensuing 10 years of intensive research and sizable funding from both government and private sources, the collective accomplishments by the DL communities are both numerous and significant. Among the tangible accomplishments, briefly stated, are:

- Advances in scientific knowledge for content creation, store, search and retrieval
- The development of scaleable testbeds for experimental research and software tools, such as metadata and standards, for managing electronic contents in a variety of knowledge domains
- Technology innovation and integration, including those that succeeded in making it to the commercial market

- Role in the development of literally thousands of operational digital libraries worldwide, far beyond the mere digital equivalents of information services provided by traditional libraries
- A hand in the transformation of traditional library science schools into academic programs that offer modern digital curricula for training future librarians and knowledge workers
- Role in the creation of many DL byproducts, including many technical publications, literally countless workshops and conferences held every year to report on the latest advances

Despite these impressive accomplishments, however, there are signs of stagnation, uncertainty or even setbacks of DL efforts. In a way, DL as a field of research is at a crossroads of two conflicting paths, one pointing to continued excitement and growth; the other leading to insignificance and even possible distinction. Nowhere is this dilemma more telling than the reports at the 2002 ASIST conference held in Philadelphia. There we heard how two equally credible DL programs faced two completely different outcomes regarding their future: one being eliminated from the university (state) funding altogether, and the other thriving on a successful operation and management, complete with a business model and financial plan. There is a lesson to be learned from this. The society (and the funding sources consequently) clearly demands more tangible evidence when it comes to continued support for research in any field of study. For the future of DL, the research community must meet the challenge of making its work sustainable over time, not only technologically but also socially, politically, and economically.

In this paper, we first discuss the roadblocks in furthering the advances in DL research and in achieving real innovations. After reviewing how other information technologies went from concepts to products and to a multi-billion business as a model for sustainable innovation, we propose a similar path for DL. A key element of this path is the integration in future DL systems of "disruptive technologies", which have the potentials of dramatically changing the conventional mindset either for new technologies or for applications.

2. Roadblocks for DL Innovation

If stagnation of progress, real or perceived, in any field of research occurs, it is important to identify where the roadblocks are in the path towards advancement or innovation. For DL research, we see several global factors that have made or are making further progress difficult.

First and foremost is the ill-formed public perception towards digital libraries. This is perhaps the most serious roadblock for DL's future. The general public by and large continues to view a digital library as the electronic version of the traditional library – where you get to use books and other materials in electronic forms either online or from the local library, for free. The broader vision for the DL circa 1994 has hardly had much effect on that outdated perception. Second, part of this ill-formed public perception is due to the overemphasis of research on the development of basic science and technology, and not enough priority given to user needs. This is reflected in what the research community do more and best, for example, technical conferences and publications. Few, if any, real large-scale open experiments and demonstrations for the benefits of the public or the industry have been done in the last ten years. The third roadblock is from a source outside the current DL arena - the commercial success in non-public domains where such e-contents and services as music and gamesare being made readily available online at lost cost. Such successes overshadow a similar need and opportunity for e-contents in the public domain. The fourth roadblock, perhaps the most important, is the lack of a business model and incentives for making public e-contents more accessible and profitable. Without such models in the DL research agenda, the path of innovation will continue to lack a sustainable force and consequently less progressive or even self-destructive.

None of these roadblocks poses a direct threat to a particular research program or project in a laboratory or to a R&D project at a company at the present time. But together, if allowed to continue, their negative impacts will have a serous consequence for DL's future.

3. The Innovation Pipeline

How do we remove these roadblocks in order for DL research to move forward and meet the new challenges of the knowledge-based society? Are there models or experiences in other areas of information technology that we can learn to jump-start the innovation process? To find answers to these questions, we take a look at the 30-year history of the innnovation pipeline for information technology.

In information technology (IT), as in many other technologies, innovations occur on the path from a new concept to a new product. This path, unfortunately, is rarely linear and predictable. Indeed, many IT innovations that we have accustomed to have followed a path with duration, complexity, and a diversity of players, institutions,

and financial support. In a 1995 report on the High Performance Computing and Communications Initiative, the Computer Science and Telecommunications Board (CSTB) of the National Research Council demonstrated this phenomenon using a number of examples in key IT technologies developed over a period of some 30 years (from 1965 – 1994). In that demonstration, which was later extended to include developments in the subsequent decade in a recent 2003 NRC report [5], the innovation pipeline is characterized by a collection of key IT technologies from their initial conception to mufti-billion dollar businesses or industries. Many of these technologies were developed in the 1960s and 1970, but nearly all of which continue to be critical to our present day economy. They include time-sharing, computer graphics, Internet, LANs, workstations, graphical user interfaces, very-large-scale integrated circuits (VLSI) design, relational databases, and reduced instruction set computing (RISC) processors. More recent innovations, such as data mining, parallel computing, redundant arrays of inexpensive disks (RAID), portable communications, world wide web, speech technology, and broadband in last mile, followed a similar path, though reaching their commercial successes much later.

In each of these technological innovations, there was invariably sustained government funded research, usually done at university labs and research centers. Such research focused on fundamental ideas and technical challenges, which would then lead to industrial research and development. A commercially successful product or products then followed. In all cases, university teams were involved from the beginning and in the transition of research into industrial quality work. The creation of a billion-dollar business is often the outcome of a sound business model, entrepreneurism, technical and community leadership, and in some instances, pure luck.

Is DL research following a similar path from conception to a successful technology? What are the chances that the future of DL research can be sustained both technologically and financially?

4. Disruptive Technologies for DL

A positive or encouraging answer to these questions may hinge on our collective abilities to steer DL research towards true innovations, departing from conventional wisdom, if necessary. What we need foremost are "disruptive" forces in emerging technologies that are not considered part of the mainstream research (in this case DL research) and yet are likely to help create drastically new advances for the field and ultimately change the field in a dramatic way. Technology historians often refer to them as the "disruptive technologies". These are the technologies – such as Marconi's wireless invention a century ago, the internal combustion engine, transistors, and the modern day Web browser – that not only led to new commercial products and industries but eventually change the world. One of the important features of disruptive technologies is that they often come from outside the mainstream. For example, the light bulb was not invented by the candle industry looking to improve its product, centuries ago. Similarly, in information technology, the concept of a Web browser for the Internet did not come from the telecommunications or network industry or even a typical computer scientist. The mainstream owners – researchers and practitioners alike – tend to focus on incremental improvements to their own research themes or products, often missing the opportunity to truly innovate.

DL's future needs to look to the disruptive technologies outside of its mainstream research. Potentially, the set of such technologies is both large and under specified. In this discussion, we'll limit ourselves to only examine a few of those that are emerging technologies and, if integrated with DL research, will most likely have a significant impact on where DL is headed in our society.

Untethered Communications. The capabilities for "untethered communications" usually refer to the union of wireless and mobile technologies. Included here are a variety of personal information appliances and wearable computing devices, which have shown extraordinary growth in the last decade. An example is cellular telephones. Nearly unheard of in 1980, they are now used by several hundred million subscribers worldwide. According to market research, the subscriber base for wireless communications services is growing 15 times faster than the wired services and in another five years, wireless and wired systems are expected to have equal numbers of users. Such public demand for untethered communications is causing intense activities in R&D, technical standards and policies both in industry and government.

What are the disruptive forces that could impact DL of the future? Developments in untethered communications will most likely help to achieve the vision of "anytime, anywhere" access to digital content, removing the walls we still have in the traditional library setting, despite a decade of DL research.

Broadband Internet The term "broadband" sometimes refers to information services other than those via Internet connectivity, which is what we consider here. Broadband Internet is a more general concept, depicting high performance for information

access and delivery over the Internet. The technology is more than that associated with particular facilities or transmission modalities used for implementation: DSL, HFC, fibre, wireless, and so on. It is really a high-performance infrastructure capable of supporting a multitude of applications and services, a means to many ends that encompass people, work, and community. Thus, for example, grid computing, which provides high end computing by connecting resources via a network, is also a key ingredient in this technology. Worldwide demand for broadband Internet has grown rapidly in recent years, with impressive penetration rates in the homes of many developed countries, especially in East Asia. By and large, up to now, broadband services have only succeeded in enabling entertainment or e-commerce applications where the contents are mostly proprietary. The next territory for broadband to conquer is the Internet's exploitation as a public space [4]. Theoretically, the technology can make digital government, education, life-long learning, personal health care, and other applications richer and more compelling and useful.

Why is Broadband Internet is disruptive technology for future DL research and innovations? Partly this is because it can provide dramatically new modalities for communication, especially within families and communities. With convergence of video, audio, text, images and so forth, the traditional library materials, plus new Internet resources, all become a digital stream that can be accessed where the high-performance network is. Old library functions can be reinforced or improved and new uses and applications, unforeseen at this point, can be invented in the future.

Distributed storage and retrieval A new technology known as "distributed storage", still in its infancy, has the potential of storing and retrieving data files in the nooks and crannies of the Internet. The idea is to free data from dependency on specific computers or systems, much in the same way books and other knowledge resources depend on traditional libraries for storage and access. The technology in its simplest form already exist – music sharing services, for example, which let people down-load and trade songs from Internet-connected PCs, are a crude distributed-storage system. Pushing this technology into including all sorts of data and media forms is one crucial aspect of this new technology. Another important feature is for such systems to be easy to use while providing security and protection of the data from being destroyed or tempered with. Theoretically, distributed storage technology would eliminate the threat of catastrophic events that make preservation of materials in traditional libraries a constant challenge.

The disruptive force behind this technology for DL is that it could fundamentally change the way digital or traditional libraries function – from acquisition, maintenance, and distribution of information resources, all the way to preservation. The power of distributed storage is yet to be explored in the research laboratories at this point, but its potential impacts on the future of DL demand our attention.

E-payment Micropayment technology has been in the market place for some time. It is associated with the concept of letting Internet users to buy digital contents and services with digital pocket change – transactions of about a dollar or less. The concept is simple and appealing in theory. Many start-up companies in the last ten years or so have tried and failed to make this technology work in the market place. Up until recently, most Internet based businesses still can't support micropayments because the processing fees from banks and credit card companies erase any profit. Then of course there is the question of demand – the question of whether consumers are willing to pay for digital contents, such as games and songs that they could get for free via the Internet. Two recent developments have changed this perspective. First, major IT players are beginning to use micropayment technology with clear success. In 2003, Apple computer announced its online music store, which sold, according to Forrester Research, more than 10 million 99-cent songs in the first four months. Second, new start-up companies are making renewed efforts to develop the next generation micropayment technologies that are easier to use, more economical, and safer. And the hope is that these new technologies will give consumers the freedom to try out new kinds of commerce on the Web, and to buy or sell an ever wider variety of digital contents.

Why is e-payment a disruptive force that could lead to dramatic innovations for DL? First, contrary to the conventional wisdom where all the contents in the libraries - digital or otherwise - should be free, micropayment technologies offer an alternative to this century old, free of charge model. By charging pennies for the bits one wants or for faster delivery of the bits, we'll be able to extend the library functions and range of information services far more beyond the traditional libraries have to offer. Selling public owned contents would be a hard sell, but it provides a business model for DL as a sustainable technology and as a financially viable public institution. A second impact is on the supply side of DLs. With micropayment on the Web, we'll be able nurture a new generation of artists and content innovators who could create, sell and trade their creative works directly to the users with financial rewards. This would be a dramatic change

from how we used to create contents, publishing books, and selling them in the market place.

The four disruptive technologies discussed above are just a few examples of a much longer list of possibilities. Nonetheless, it is worth summarizing their key features and what their potential impacts might be on the future of DLs and to the knowledge-based society. This summary is given in Table I below.

5. A Transformational Research Agenda

If we add up the key points discussed above, what we come up with is a proposal for a new research agenda capable of transforming DLs into a sustainable technology and permanent institution. This agenda includes the following three key elements:

- Link up with disruptive technologies: Future content – and media-based research must consider integrating these technologies into its research paths as a basis for innovations

- Focus on "DL middleware": Government and private research programs should place increased emphases on standards and software tools that serve the foundation of DL "middleware" across disparate domains of knowledge, computing platforms, and communications channels.

- Create a Business Model: Public e-contents can be and need to be made vastly more accessible by following a technologically and financially sustainable path.

Table 2 below depicts a possible scenario where we can add DL middleware, for lack of a better generic title, as a new entry into the IT innovation pipeline. The milestones of past DL research and the projected future research and development leading up to the birth of a billion-dollar business are merely suggestive. Turning this scenario into reality is the challenge for the entire DL community.

6. Concluding Remarks

The first decade of DL research invokes two reactions. The first calls for a self-assessment of accomplishments by the research community, which by any measure are impressive and successful. A second reaction usually comes from the public and those who have provided funding and other resources for the research. It is the second and outside reaction we must discern carefully and use it to guide our future work. This paper has proposed a way to develop the next research agenda, not to incrementally improve what we have done before but to push the envelope of what is perceived possible in relation to other technologies, models of success, and to the society as a whole.

References

1. Randy Katz and Y.T. Chien, Eds., "Information Infrastructure Technology and Applications", Report by the IITA Interagency Task Group, HPCC National Coordination Office, February 1994.
2. Y.T. Chien, "The role of digital libraries in the transformation of society: a perspective from the United States", Proceedings of the 23rd Digital Library Workshop, University of Tsukuba, Tsukuba, Japan, November 2002.
3. Y.T. Chien, "Whither Digital Library Research? The Case of a "billion-dollar" Business, Presentation at the Research Center for Knowledge Community Colloquium, University of Tsukuba, Tsukuba, Japan, October 31`, 2002.
4. Jean Camp and Y.T. Chien, "The Internet as Public Space: Concepts, Issues, and Implications in Public Policy", in: *Readings in CyberEthics* (Eds. R. Spinello and H. Tavani), Jones and Bartlett Publishers, 2001; Also published in ACM Publication of Special Interest Group on Computers and Society. September 2000.
5. Report of the National Research Council, Innovation in Information technology, National Academy Press, 2003.
6. Report of the National Research Council, The evolution of untethered communications, National Academy Press, 19997.
7. Report of the National Research Council, Broadband – bringing home the bits, National Academy Press, 2002.
8. NSF CISE Report on Revolutionizing science and engineering through Cyberinfrastructure, January 2003.
9. Clayton M. Christensen, *The Innovator's Dilemma* – When new technologies causes firms to fail, Harvard Business Press, 1997.
10. Gregory Huang, "The Web's new currency", MIT Technology Review, December 2003/January 2004.
11. Wade Roush, "Computers that speak your language", MIT Technology Review, June 2003.

sTable 1. Contribution of Disruptive technologies to DL's Future

Technology	***Components or Functionalities***	***Potential Innovations and Impacts on Society (examples)***
Untethered communications and Platforms	• Wireless and mobile • Wearable computing • Satellite communications • Portable Information appliances	• Access and assistive tools for the disabled • Anytime, anywhere DL access • Personalized computing platforms
Broadband Internet	• High speed networks • Grid computation • Embedded computers	• Streaming of large, high quality audio-video and date files • DL access for home or distance education, life-long learning • Computer-mediated communications
Distributed Storage and Retrieval	• Freely moving and replicating digital files across machines • Making files easier to access and maintain while eliminating the threat of damage	• DL Middleware track distributed data files • Software and programming systems for DL preservation • Personalized computing environments anywhere, anytime
E-payment	• Charging customers for using e-content without incurring costly transaction fees	• Micropayments for public e-content in exchange for vastly improved information services • New generation of e-content suppliers: artists, writers, etc. motivated by the economic potentials of e-payment

Table 2. Adding a new entry to the Innovation Pipeline – DL middleware