Introduction

Goal of this conference is to bring together experts in Science, Technical, and Medical (STM) journal publishing, both producers and users of these publications to:

1. Identify the recent technical changes in publishing, and other factors, that influence the decisions of journal publishers to produce journals electronically.

2. Identify the needs of the STM community as users of journals, whether electronic or printed.

3. Discuss the responses of the not-for-profit and commercial STM publishers and of other stakeholders in the STM community to the opportunities and challenges posed by the shift to electronic publishing.

4. Examine the spectrum of proposals that has been put forth to respond to the needs of users as the publishing industry shifts to electronic information production and dissemination.

Although a major aspect is to examine the various business models for making the scientific literature (and related information products and services) available, it is also important to consider the implications from the various e-publishing developments on the scientific enterprise:

1. How is EP affecting the practice of scientific research;
2. The communication of research results to scholars and others (the public?)
3. The curation of data, evaluation of research, and archiving of results?

Ultimate purpose: identify the main issues and problems that the STM community needs to confront and resolve if it is to exploit the opportunities and cope with the challenges presented by this very rapidly evolving technology.

When Ted Shortliffe asked me to kick-off this discussion, I was a bit perplexed.

As a one-time (and now once-again) scientist, and a has-been university administrator, my basic approach to such issues has been one of avoidance:

Of the hassles of publishing as a scholar

And of the costs of acquisition and maintenance of STM journals as a university provost and president

But, after some further (desperate) thought, I found a couple of potential hyperlinks to the subject from my personal experience

As a voracious producer and consumer
Except now I write books rather than papers
But I also dabble in multimedia and 3D simulations
And, at last count, I have half-a-dozen computers and over a terabyte of data sitting on my desktop on Firewire disk drives (God only knows what is contained in them)

During my last years as president, I built a major "library of the future" at Michigan we call the Media Union (which also happens to be where some of the servers for JSTOR are housed)

A Rather Unique experience: Supreme Court trial on diversity
Entire electronic inventory now online (3,000 or so documents!!!)

Chair of ITFRU and IT Forum

As I recall, in "the Good Old Days":

Scientific communication
= Xerox + mail = preprints (to the "in crowd")

Journal publications?
Too late (1-2 year delay…already "archives"

Too expensive (pages charges limited publication)
Too inaccurate (transition to computer composition)
What mattered? Tenure?
Based on peer-evaluated quality of work
(as much through presentation as publication).

What has changed?

Scientific communication = Internet/websites = still preprints

"Document" now a collaborative platform

Journal publications?
Still too late, too expensive
(and few of the "in crowd" read them anyway…
mostly for graduate students)

Tenure?
Based on quantitative measures (number, citations)
Grantsmanship

Another perspective: As a university provost and president

Internet-1 (NSFnet) and now Internet2

Staggering cost of scholarly journals (and libraries)

Big Ten effort to create a single library

Mellon – JSTOR (run servers from UM)

The Current Situation: A chaos of concerns…and responses

1) The disturbing trends in STM publishing seem to be continuing:

2) Access to scientific, technological, and medical (STM) information is increasingly expensive and restricted.

3) The amount of STM information generated at research institutions continues to grow, and so have the number of STM journals

4) Journal subscription prices continue to escalate, both for print and electronic

5) University library budgets fail to keep pace with the increase in research expenditures and subscription rates, resulting in subscription cancellations and reduced acquisitions

A Case Study: UM

Price inflation on EP has continued to run well ahead of the CPI. The total increase over ten years could be 75-100% but Jim should search this a bit in the literature for confirmation. We know that the annual rate has often been in the 8% to 12% range any given year, sometimes higher for individual publishers.

Another group of materials where we have seen a sharp difference in pricing is in the area of the reference tools. The increase in annual price on such materials has gone up significantly, sometimes as much as 600% over the print cost of the bound volume. The publisher argues that the added value and distributed access that the electronic version provides warrants the increased price. And there is no alternative as each item is often a unique resource.

Second, we're not just dealing with print but have had to accommodate a variety of pricing schemes designed by the publishers to protect print price revenues and then recoup their investments in converting content to a digital form and building systems to distribute these data. On average, we probably pay 15% more when we add electronic access to a title we're already paying for in print. {That is, the two formats combined amount to about 115% of the previous print subscription alone.

Third, the nature of searching full text across large numbers of titles has led us to add titles by subscribing to the full stable of titles offered by particular publishers ("supersizing").

As noted in the preample to this conference, the now universal access to and use of the Internet and other digital information technologies by the STM research community has transformed many aspects of the research process, not least the ability to communicate data, ideas, results more quickly, broadly, and openly than was possible through traditional print publications in the past.

New technologies have created fundamental changes in the production, management, dissemination, and use of all types of information. Researchers are now able to make available independently their data and articles online where the information may be easily found, browsed, annotated, critiqued, downloaded, and freely shared. This is resulting in significant changes to the linear path of writing, refereeing, and reviewing of publications, since all of these functions can now be performed concurrently.

1) For publishers and authors the question is: "How many copies of the work will be sold (or licensed) if networks make possible planet-wide access? Their nightmare is that the number is one. How many books (or movies, etc.) will be created and published online if the entire market can be extinguished by the sale of the first electronic copy?"

2) The nightmare of consumers is that the attempt to preserve the marketplace leads to technical and legal protections that sharply reduce access to society's intellectual and cultural heritage, the resource that Jefferson saw as crucial to democracy.

Reactions and counter reactions

The first and most obvious reaction to the high cost of STM journals has been subscription cancellation—initially out of necessity, but also as a strategy. Cancellations, however, drive up subscription charges as publishers realize less revenue and try to find other ways of regaining it.

Libraries often find themselves at disadvantage negotiating subscription licenses singly, so they began forming consortia for the purposes of negotiating subscription licenses and packages.

There are countervailing forces acting to suppress prices such as SPARC, of which Michigan Library is a member. {A unit within the Association of Research Libraries, it is working to find alternative publishing channels for scholarly information. SPARC is The Scholarly Publishing & Academic Resources Coalition, which now has global membership.

Within the CIC, we have begun a program where a list of a publisher's electronic titles is reviewed by the thirteen libraries. Each institution selects titles that are particularly important for their local programs and agrees to continue the paper subscription, ensuring one copy available within the CIC. Other members of CIC can then comfortably cancel their print subscription if not needed locally, relying on the electronic, but knowing there is a print backup within the CIC to which they have ready access.

In several instances, editorial boards have protested against the steep increase in journal prices, resigning from commercial publishers and going to less expensive publishers (including scientific societies)

Instead of ownership resulting from first-sale doctrine—once a paper publication is paid for, the buyer owns it, can lend it, can sell it, and can reread it without additional payment—electronic STM information is licensed for a specified length of time, for a specified number of users, and for limited access. The purchaser buys access to the information but does not own it. Institutional subscription rate structures for electronic access and print copies are complicated, and libraries often negotiate rates with publishers, either singly or as consortia.

Lots of experiments:

SPARC is one of several otheruniversity led publishing initiatives with better pricing and terms-- e.g., Highwire {Stanford}, Allen Press, Scholarly Publishing Office {Michigan}, Euclid {Cornell}, etc. There is also the Institutional Repository movement that may challenge commercial publishing's baser instincts about pricing.

Variants on University Microfilms (e.g., creating sustainable financial models for publishing an "edition of one"), e.g. Cornell Experiment: We seek to have the decision to publish based upon intellectual merits without regard to sales potential. Being "published" by a university press could be redefined to include online distribution, when supplemented with on-demand, printed and bound copies. Interested readers could read the book online, or elect (at their expense) to order a single, perfect bound copy. This would preserve the convenience of the traditional process, making bound copies available as needed, without causing the press to be exposed to the

financial risk of a large inventory of copies that may or may not be sold.

If the faculty member's dean agreed upfront to fund the direct costs for a new faculty member's first book to be published in this manner, the press would be relieved of its financial risk for the book. At the same time the faculty member would be reassured that the dean's funding represents an explicit affirmation of the legitimacy of this publishing approach, and that such a digital publication would not disadvantage him/her in tenure and promotion considerations. (A bit analogous to lab equipment startup funds.)

Open Source Strategies

The university has historically be about open inquiry and communication although the faculty routinely passes off their intellectual contributions to commercial publishers who add value and then disseminate the products in exchange for payments under copyrights and licenses. The success of the open software moment that has given the world Linux, the Apache web server are encouraging exploration of alternative sustainable models of dissemination of intellectual properties including both content as well as the source code for software-based tools and services. Such open source communities are interesting examples of a cooperative meritocracy and is now being proposed for other types of intellectual property (Open Archives Initiative, Open Knowledge Initiative, the Creative Commons, the Comprehensive Collaborative Framework, and the MIT Open Courseware project.

Some contend that STM journal articles can be made available for free online and that revenue for such efforts can be gained without charging subscription rates: the cost of publishing online is less than publishing on paper, and the means to recover costs can include formulating value-added products (e.g., reprints with extra information). The principal means of making revenue would be to "charge for disseminating articles rather than for accessing them"—for instante, by building the charges for dissemination into the research grants that generates the information.

This is consistent with the traditions and values of academia and it offers a potentially important counterbalance to the aggressive digital rights management movement of the commercial entertainment and media worlds. It may also run against the pressures for privatization and commercialization of data and information resulting from publicly supported research.

This also reinforces the definition of university as a public good rather than a market commodity. (Something we worry about very much these days in higher education.)

So Why All the Bother?

In summary:

1) Advances in technology have produced radical shifts in the ability to reproduce, distribute, control, and publish information.

2) With its commercialization and integration into everyday life, digital information infrastructure has run headlong into existing publication practices, policies, and laws.

Why are the issues so difficult?

1) The stakeholders are many and varied.

2) Content creators have different agendas, handle IP according to varying strategies, and look for different kinds of return on their investments.

3) Fundamental legal concepts can be interpreted differently.

4) The economics of information products and IP can be subtle.

More specifically:

> Funders provide money for research, and authors produce and write up results of research. Scientific, technical, and medical research is funded by organizations that wish to see STM

information advanced and made available to the public. That content—a public good—is provided by authors to publishers free of charge. Editors and peer reviewers evaluate articles for publishers, also free of charge.

Although authors, editors, and reviewers do not charge for their labor and may be motivated to contribute to the public good, they are not without compensation: in a publish-or-perish atmosphere, scholars give away content in order to gain tenure and promotion. If they publish in, or work on, the "right" journals, they gain prestige.

Publishers are intermediaries that do not pay for STM content but make it available in published form. They can be profit or not-for-profit; both are paid. Income may be designated profit or used to cover overhead. It may be used to pay for less successful in-house ventures.

Libraries are also intermediaries that provide access to STM content. They pay for content but do not usually charge for providing access to it. Therefore libraries, not the users of the library journals—for the most part, and so far—are affected by subscription cost changes. Given the turnaround from first-sale ownership of paper subscriptions to licensing access for electronic subscriptions, libraries now have the role of limiting access according to licensing agreements. These developments put libraries in the position of paying for access to increasingly expensive STM information and for policing it.

Content users are readers of STM content. They may either pay for content or access it for free through libraries. Although many content users are academic scholars and scientists in universities and government, most are in industry and rely heavily on not-for-profit-sponsored research publications.

So we have a dilemma, perhaps best stated by John Perry Barlow in an article in Wired magazine several years ago:

Barlow's Enigma: "If our property can be infinitely reproduced and instantaneously distributed all over the planet without cost, without

our knowledge, and without its even leaving our possession, how can we protect it? How are we going to get paid for the work we do with our minds? And, if we can't get paid, what will assure the continued creation and distribution of such work?"

"Since we don't have a solution to what is a profoundly new kind of challenge, and are apparently unable to delay the galloping digitization of everything not obstinately physical, we are sailing into the future on a sinking ship. This vessel, the accumulated cannon of copyright and patent law, was developed to convey forms and methods of expression entirely different from the vaporous cargo it is now being asked to cary. It is leaking as much from within as from without."

Michael Crichton speculated even further in a Wired article entitled "Mediasaurus". Today STM articles, journals, books, archives, and databases are indistinguishable parts of the content industry—they are all made of bits—and organizational, technological, and legal treatment of different kinds of bit structures are starting to look similar.

"To my mind, it is likely that what we now understand as the mass media will be gone within 10 years. Vanished without a trace. The media are an industry, and their product is information. The media have always been driven by technology, but it is surprising how many of their attitudes and terminologies are very old. (Steroeotype and cliché are 18th century printers' terms that refer to metal type.) The key to the Mediasaurus is the coming end of media's information monopoly. Soon we will have AI based agents roaming the databases, downloading stuff that I am interested in, and assembly for me a front page or a nightly news show that addresses my interests." (Is Google.News a precursor of this trend?)

So how can we approach these issues. Your conference approaches this in one way, organized along 6 panels that I will come back to review briefly.

But with more limited time, I would like to consider the issues from only three perspectives:

1. The Evolving Practice of STM Research

2. Policy and Legal Issues

3. The Exponential Evolution of Digital Technology

First, a couple of basic observations, however:

The Age of Knowledge

Looking back over history, one can identify certain periods of profound change in the nature, the fabric, of our civilization such as the Renaissance, the Age of Discovery, and the Industrial Revolution. There are many who contend that our society is once again undergoing such a fundamental shift in perspective and structure.

The signs are all around us. We are evolving rapidly into a postindustrial, knowledge-based society, just as a century ago an agrarian America evolved into an industrial nation.1 Today industrial production is steadily shifting from material- and labor-intensive products and processes to knowledge-intensive products. A radically new system for creating wealth has evolved that depends upon the creation and application of new knowledge.

We are in a transition period where intellectual capital, brainpower, is replacing financial and physical capital as the key to our strength, prosperity, and well being. In a very real sense, we are entering a new age, an *age of knowledge*, in which the key strategic resource necessary for prosperity has become knowledge itself, that is, educated people and their ideas.2 As our society becomes ever more knowledge-intensive, it becomes ever more dependent upon those social institutions that create knowledge, that educate people, and that provide them with knowledge and learning resources throughout their lives.

Unlike natural resources such iron and oil that have driven earlier economic transformations, knowledge is inexhaustible. The more it is used, the more it multiplies and expands. But knowledge is not available to all. It can be absorbed and applied only by the educated

mind. Hence, schools in general and universities in particular will play increasingly important roles as our society enters this new age.

If, as is often claimed, societies are seeing a shift in economies as significant as the industrial revolution, with the transition to knowledge and information as a major source of wealth, then intellectual property may well be the most important asset in the coming decades.

As the Panel 4 summary statement "The word *publication* is rooted in *the act of making public* and since about 1450 has meant doing so largely by composing a master, printing with ink on paper, manufacturing books and journals, and distributing them by physical transportation. Bringing a publication into use by an individual, as opposed to providing only potential access, is a combination of push and pull between consumer and producer – marketing, assignment, referral, browsing and searching. This traditional process of publishing, as well as the nature of the objects produced by the publishing process, is changing radically in the electro-optical digital age."

In a physical world publication has three important characteristics: It is public, it is irrevocable, and it provides a fixed copy of the work. In the digital world, none of these may be true. The information infrastructure blurs the distinction between publication and private distribution.

Returning again to John Perry Barlow:

"Notions of property, value, ownership, and the nature of wealth itself are changing more fundamentally than at any time since the Sumerians first poked cuneiform into wet clay and called it stored grain. Humanity now seems bent on creating a world economy based on goods that take no material form. In doing so, we may be eliminating any predictable connection between creators and a fair reward for the utility or pleasure others may find in their works. Since it is now possible to convey ideas from one mind to another without ever making them physical, we are now claiming to own ideas themselves and not merely their expression. And since it is likewise now possible to create useful tools that never take physical

form, we have taken to patenting abstractions, sequences of virtual events, and mathematical formulae—the most unreal estate imaginable. Only a very few people are aware of the enormity of this shift, and fewer of them are lawyers or public officials. In certain areas, this leaves rights of ownership in such an ambiguous condition that property again adheres to those who can muster the largest army–except that this time the armies consist of lawyers.

In our efforts to extend these policies to electronic expression, we run the risk of placing in peril the ultimate source of intellectual property—the free exchange of ideas. The greatest constraint on future liberties may come not from government but from corporate legal departments laboring to protect by force what can no longer be protected by practical efficiency or general social consent.

Back to the Three Key Perspectives

1. The Changing Nature of STM Research

Mention and show NSF Cyberinfrastructure Report

The nature of knowledge work is changing, increasingly based upon cyberinfrastructures such as collaboratories or grid communities. In the sciences there is a trend towards more collaborative, multi-disciplinary work. The process of knowledge creation – experimentation, analysis, theory development, forming conclusions, is occurring in virtual organizations enabled by cyberinfrastructure and thus making intermediate as well as the final products of the work more widely available faster and sooner. The net effect is movement to a more continuous-flow model of scholarly communication rather than more discrete, batch processing. Examples include institutional respositories, digital asset management systems, pre-print servers, and electronic journals. Digital technology has opened and disaggregated the traditional print-on-paper publishing chain and academic libraries now have the opportunity to add value at various parts of the process.

Distributed network computing technology provides an infrastrcture for federating people, information, computational tools and services, and specialized facilities into virtual organizations for knowledge

based activities at multiple levels of aggregation–team, firm, university, etc. These new environments included collaboratories, grid communities, and e-science environments.

A new age has dawned in S&E research, pushed by continuing progress in computing, information, and communication technology, and pulled by the expanding complexity, scope, and scale of today's challenges. The capacity of this technology has crossed thresholds that now make possible a comprehensive "cyberinfrastructure" on which to build new types of scientific and engineering knowledge environments and organizations and to pursue research in new ways and with increased efficiency.

The emerging vision is to use cyberinfrastructure to build more ubiquitous, comprehensive digital environments that become interactive and functionally complete for research communities in terms of people, data, information, tools, and instruments and that operate at unprecedented levels of computational, storage, and data transfer capacity.

Electronic Publishing

Scholarly publication is also changing rapidly. Most scientific results now appear first as preprints on the web, perhaps authored by a collaborative of scientists. Only many months (or years) later will appear in refereed journals, then more likely to go immediately into the dusty stacks of a library than to be read as a physical document.

A half century ago Vannevar Bush wrote “our methods of transmitting and reviewing the results of research are generations old and by now are totally inadequate for their purpose”. And yet much remains the same today, except that the volume of literature has increased vastly with prices soaring to staggering levels driven by a monopolistic publishing industry.

There is a recursive relationship between information technology and scholarly communication, since rapid advances in each depend upon the other. Robert Lucky, former president of Bell Laboratories, suggests that “in their influence on how science is transacted, the Internet and World Wide Web have had the greatest impact of any

communications medium since possibly the printing press."3 As with learning, these new electronic media allow the formation of spontaneous communities of unacquainted users, linked together in the many-to-many topology of computer networks. Researchers can now follow the work in their specialization on a day-by-day basis through web sites. As Lucky notes, "Who could have dreamed even a decade ago that we would have instant access to a billion documents from the comfort and privacy of our office or laptop?"

Yet even today, science is still characterized by the publication of research results. The current confusion between traditional scholarly publication, in established journals characterized by peer review–and high costs–and less formal Net-based communications, linking together scholars essentially instantaneously, continues to present a challenge. But here too technology is evolving, with the rapidly evolving use of Web sites that serve as portals to integrate material of interest to particular scholarly disciplines.

As Lucky predicts, "we are headed technologically to a time when bandwidth and processing speed will be unlimited and free. Latency will approach the speed-of-light delay. Service quality will approach the "five nines" (99.999%) availability of the traditional telephone network. And encryption will finally be freed of its political restraints, assuring security and privacy."

Electronic publication opens vast possibilities, not the least of which is freedom from the monopoly and pricing of commercial publishers. Yet, important standards of scholarly publication such as critical review and attribution of credit must be retained even as digital technology reshapes scholarly communication. So too, the access to scientific instrumentation and data and the rapid dissemination of research results allowed by digital media raises complex issues of investigator control and sharing in scholarly investigations. The rising concerns of researchers with protecting intellectual property because of its potential commercial value can also interfere with teamwork and sharing.

The reality is that electronic publishing will become the dominant mechanism for both reading and publishing scholarly materials by the end of this decade. Yet, to achieve this goal, the scholarly

community will demand a new paradigm for scholarly communication capable of providing open online access to the work of scholars without payment, online repositories of high quality materials, and a stable economic model to sustain these resources.

The Library

The preservation of knowledge is one of the most rapidly changing activities of the university. The computer--or more precisely, the "digital convergence" of various media from print-to-graphics-to-sound-to-sensory experiences through virtual reality--will likely move beyond the printing press in its impact on knowledge. Throughout the centuries, the intellectual focal point of the university has been its library, its collection of written works which preserve the knowledge of civilization. Today such knowledge exists in many forms--as text, graphics, sound, algorithms, and virtual reality simulations--and it exists almost literally in the ether, distributed in digital representations over worldwide networks, accessible by anyone, and certainly not the prerogative of the privileged few in academe.

This poses a particular challenge to the library, shifting it from a focus on collecting and archiving knowledge resources (most commonly in written form) to assisting scholars to navigate through a vast array of digital knowledge resources scattered through cyberspace. Just ask colleagues when they last visited a library. Probably years ago. The reason is simple: The library is no longer a place. It is a utility. It is becoming less a collection house and more a center for knowledge navigation, a facilitator of information retrieval and dissemination.4 The campus library has become less central to most researchers' lives, with digital telecommunications allowing access to primary and secondary materials online. Libraries no longer chain their books to the wall; indeed, they no longer require physical visits.

Of course, books are also changing rapidly. We are already beginning to see the early "e-books"–really small laptop computers optimized to display text. New publishing paradigms are appearing, such as the use of the Net to provide downloads for e-books. But this technology will evolve very rapidly, with the development of "electronic ink"

that allow electronic books to closely resemble conventional books, albeit with the powerful tools of network access to the libraries of the world, hypertext or web-links, the display of non-verbal materials such as video and virtual, and eventually the use of artificial intelligence and software agents. In fact, the key new feature of e-books will be their use of interactive multimedia as a communications medium and their instantaneous access to vast knowledge networks.

Furthermore creating and managing new forms of archival and support material has become increasingly important to many research fields. Databases of longitudinal economic data, archives of all documents in a particular language or from a particular period, digital video archives of survey interviews, raw data from observatories and nuclear accelerators are increasingly the source material for research, and the management of such vast data sources are stimulating new methods of scholarship. Furthermore, such data archives are typically not physically localized nor controlled by a single institution.

As familiar as the exponential growth in computing, storage, and networking power is the exponential growth in digital information and data. Most all scientific and technical literature is now created in digital form, and large quantities have been converted to digital retrospectively. But pricing and terms and conductions for use continue as major issues, since they force academic libraries to collect a smaller and smaller fraction of the overall literature. The primary access to the latest findings in a growing number of fields is through the Web, then later through classic preprints and conferences, and only afte that through refereed archival papers. The traditional linear batch processing approach to scholarly communication is changing to a process of continuous refinement as scholars write, review, annotate, and revise in near-real time using the Internet. Through data mining of massive digital archives new knowledge is being discovered in problems areas never intended at the time of the original data acquisition.

It is inevitable that digital genres will dominate and that paper as the prim ary form for access will be secondary. Most documents are now born digital and there will also be continuing growth in new

multimedia genres that have no print or paper equivalent. We are on the path to digital as the primary content of libraries and paper as a secondary augmentation, largely as the format of last resort or if the digital version or adequate digital metadata is not available. Advocate a flip in perspective of library from "a collection of physical, printed resources and associated services augmented by digital resources and services" to "a library is a federated collection of (physically distributed) digital resources and services augmented by printed resources and associated services.

2. Legal and policy issues

Thoughout the history of copyrights and patents, the proprietary assertions of thinkers have been focused not on their ideas but on the expression of those ideas. The ideas themselves, as well as facts about the phenomena of the world, were considered to be the collective property of humanity. Law protected expression, and with few (and recent) exceptions, to express was to make physical.

The information infrastructure promises more for intellectual property–more quantity, quality, and access–while imperiling one means of rewarding those who create and publish it. It has the potential to demolish a careful balancing of public good and private interest that has emerged from the evolution of U.S. intellectual property law over the past 200 years. The public good is the betterment of society that results from the constitutional mandate to promote the "progress of science and the useful arts"; the private interest is served by the time limited monopoly (a copyright or patent) given to one who has made a contribution to that progress. The evolving information infrastrcture presents a leap in technology that may well upset the current balance, forcing a rethinking of many of the fundamental premises and practices associated with intellectual property.

Concern that the current legal structure of copyright may set up roadblocks against the fulfillment of its promise.

"To hear them tell it, copyright is a law invented by publishers solely to serve their own financial interests, a personal use exemption to copyright law exists for the convenience of scholars, and any

educational use of copy righted material is, by definition, a fair use. And the most pernicious of all : Copyright and intellectual freedom are fundamentally opposed, locked, like good and evil, in a Manichean struggle for the soul of the university" (Peter Givler, Ex Dir of Am Assoc of University Presses)"

From the publishers point of view, copyright is the bedrock, the legal foundation of the business. But they tend to confuse the STM community with dense, technical rules about how to comply with the law without ever explaining why anyone should bother. Hollywood has reinforced this Big Brotherish air with the term extension for Mickey Mouse.

What is a copyright? An idea in a copyrighted work isn't patented; it isn't even copyrighted. As the law itself says, "In no case does copyright protection for an original work of authorship extend to any idea, procedure, process, system, method of operation, concept, principle, or discovery, regardless of the form in which it is described, explained, illustrate, or embodied in such work". COPYRIGHT PROTECTS FORMS OF EXPRESSION, NOT THE IDEAS OR FACTS BEING EXPRESSED.

Words are things. Copyright is a specialized form of property law that recognizes that works of original expression belong to the person who created them. It gives authors legal control of their own texts, creating a system for maintaining textural integrity. Second, it created our modern sense of what the profession of author means: namely to be someone whose reputation rests on being recognized as the creator of the precise texts published under his or her name.

Copyright grants the creator of original written work the exclusive right to publish it for a certain length of time, and makes that right transferable to someone else. It creates the basic legal mechanism that allows publishing costs to be recovered from the marketplace. A business investment made in publishing a work protected by copyright can be recovered from sales. By shifting the financing of publication from patronage to the marketplace, copyright laid the foundation for the enormously expanded range of ideas and information published today.

The Supreme Court (Ruth Bader Ginsburg) recently affirmed the important relation between free speech and copyright: "Indeed, copyright's purpose is to promote the creation and publication of free expression. As Harper & Row observed: 'The Framers intended copyright itself to be the engine of free expression. By establishing a marketable right to the use of one's expression, copy-right supplies the economic incentive to create and disseminate ideas.'"

Again, remember Barlow: "We are sailing into the future on a sinking ship. This vessel, the accumulated cannon of copyright and patent law, was developed to convey forms and methods of expression entirely different from the vaporous cargo it is now being asked to cary. It is leaking as much from within as from without."

3. The Evolution of Digital Technology

Two years ago the presidents of our National Academies launched a project to understand better the implications of information technology for the future of the research university, which I was asked to chair.

In the United States, the National Academies (i.e., the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine) have a unique mandate to monitor and sustain the health of the nation's research universities as key elements of the national research enterprise and the source of the next generation of scientists, engineers, and other knowledge professionals.

This role becomes particularly important during periods of rapid change. It was from this perspective that last year the presidents of our National Academies launched a project to understand better the implications of information technology for the future of the research university.5

The premise of the National Academies study was a simple one:

> The rapid evolution of digital technology will present many challenges and opportunities to higher education in general

> and the research university in particular. Yet there is a sense that many of the most significant issues are neither well recognized nor understood either by leaders of our universities or those who support and depend upon their activities.

The first phase of the project was aimed at addressing three sets of issues:

1. To identify those technologies likely to evolve in the near term (a decade or less) which could have major impact on the research university.

2. To examine the possible implications of these technology scenarios for the research university: its activities (teaching, research, service, outreach); its organization, structure, management, and financing; and the impact on the broader higher education enterprise and the environment in which it functions.

3. To determine what role, if any, there was for our federal government and other stakeholders in the development of policies, programs, and investments to protect the valuable role and contributions of the research university during this period of change.

Our steering group has met on numerous occasions to consider these issues, including site visits to major technology laboratories such as Bell Labs and IBM Research Labs and drawing upon the expertise of the National Academy complex and last year we pulled together over 100 leaders from higher education, the IT industry, and the federal government, and several private foundations for a two-day workshop at the National Academy of Sciences to focus our discussion.

(Show IFTRU Report)

Let me mention the key conclusions from first phase of this study:

Point 1: The extraordinary evolutionary pace of information technology will not only continue for the foreseeable future, but it could well accelerate on a superexponential slope.

Hardware Technology Trends

- Processing (Moore's Law) (increasing 40% per year)
  - Current speed record: 150 GHz chips
  - Japan Earth Simulator at 40 TeraFlops
  - (ASCI Purple: 100 TFlops: LLNL 360 Tflops; Blue Gene at 1 PFLOP)
- Disk storage (increasing 60% to 100% per year)
  - 3.5 disk can hold 320 Gb
  - Far cheaper than paper or microfilm
- Bandwidth
  - Lab demo on single fiber: 11 Tb/s
  - Real communication at 40 Gb/s
- Mobility
  - 802.11 (a, b, g, I) at 55 Mb/s and beyond
- Displays
  - Full wall projections
  - Resolution much better than paper

Technology Directions (Today, 2003-2006, 2010)

- Access Bandwidth: 56 kb/s -> Mb/s -> 100 Mb/s-1 Gb/s
- Backbone Bandwidth: 155 Mb/s -> Tb/s -> Pb/s
- Intercontinental Bandwidth: 45 Mb/s -> 3 Tb/s -> many Tb/s
- Wireless: 32 kB/s -> 55 Mb/s -> Gb/s
- Enterprise database: 30 TB -> PB -> 10 PB +
- Supercomputing: 40 TFLOPS -> PFLOPS -> 100 PFLOPS
- Display: .5 Mpixel, 5. sqft -> 9 Mpixel, 60 sqft > much more

Software and System Technology

- Algorithm improvements
- Embodiment of techniques and processes into software
- Formalization and standardization
- People are the exception rather than the main line
- Distribution of computing, data, applications, and services

- Grid intercollection of resources
- Services as unit of IT, rather than bare-bones data and processing
- Semantic Web (machine-readable descriptions)

Point 2: The impact of information technology on the university will likely be *profound, rapid, and discontinuous*–just as it has been and will continue to be for the economy, our society, and our social institutions (e.g., corporations, governments, and learning institutions).

Information and communications technology will affect the activities of the university (teaching, research, outreach), its organization (academic structure, faculty culture, financing and management), and the broader higher education enterprise. However, at least for the near term, meaning a decade or less, we believe the research university will continue to exist in much its present form, although meeting the challenge of emerging competitors in the marketplace will demand significant changes in how we teach, how we conduct scholarship, and how our institutions are financed.

Universities must anticipate these forces, develop appropriate strategies, and make adequate investments if they are to prosper during this period. Procrastination and inaction are the most dangerous courses for universities during a time of rapid technological change.

Point 3: It is our belief that universities should begin the development of their strategies for technology-driven change with a firm understanding of those key values, missions, and roles that should be protected and preserved during a time of transformation.

The university will have to adapt itself to a radically changing world while protecting is most important values and traditions such as academic freedom, a rational spirit of inquiry, and liberal learning.

Although we can predict the rapid pace of change of the technology, it is far more difficult to predict the impact on human behavior and social institutions with any precision. Nevertheless, higher

education must develop mechanisms to at least sense the potential changes and to aid in the understanding of where the technology may drive it.

It is therefore important that university strategies include: 1) the development of sufficient inhouse expertise to track the technological trend and assess various courses of action; 2) the opportunity for experimentation, 3) the ability to form alliances with other academic institutions as well as with for profit and government organizations.

Organization of Conference

Panel 1: Costs of Publication

Some initial questions include:

What does STM publishing cost, and what are the dominant cost components (direct and indirect)? (e.g., submittal, pre-screening, reviewing, revising, re-reviewing, production preparation, and eventual production for online and in-print dissemination; merging of article contents to online services)

What costs are rising most rapidly?

Why does it cost so much, and how might it be made less costly? How big are the variations in cost from one discipline to another, from one medium to another and what causes these differences? It is important to note that the mega-journals have a much different cost structure than medium and small journals, as well as commercial and not-for-profit.

What we want to know from our speakers is:
What's currently happening in their business model? Does it work? Why/why not? Will it work long-term? Why/why not? Is it what we need to have happening now and later? What might we need to start doing differently that will get us to a better outcome for the scientific community?

Panel 2: Publication Business Models and Revenue

The core question is: How to organize the delivery and rights management across modes and access points so as to recover production costs, induce ongoing investment and development, and maximize social welfare?

Background considerations include the following:

Who values STM publications? (e.g., teachers, researchers in various sectors, students in higher education, practitioners, general public, and, of course, authors).

What are the access points? Developed and developing world; various institutions, such as higher education, government and industry labs, STM workplaces (e.g., hospitals, engineering design facilities, etc.); personal workstations.
Who are the access providers? Authors themselves, publishers, aggregators.

What are the access modes? Bibliographical, abstract, single view, multiple / unlimited views, print and file, local electronic copy, server copy

Discussion questions:

Both authors and readers benefit: Should revenue extraction be from both? What are the consequences for transaction costs and creation / consumption efficiency (social welfare)?

There is a public interest in dissemination of knowledge in addition to its creation. To what extent should public institutions (government, NGOs, universities) pay for dissemination? To whom should publicly-funded dissemination be available?

There is value both in the content, and in the publication of the content. How well do markets work in extracting the different values and directing them to the different producers so that we get efficient investment in content and in publication?

What are the implications for science from the different options?

Panel 3: Legal Issues in Production, Dissemination, and Use

The discussion in this session is divided into three areas:

(1) Copyright basics: ownership and rights. Before addressing the issues surrounding the grant of copyright from authors to publishers, it is important to review the "upstream" question of whether academic and research authors are initial copyright holders, or whether the institutions that employ them in fact hold the copyrights. The initial position matters for a variety of reasons, which need to be better understood. With regard to rights: What rights does copyright confer on authors (including the right to make available for free)? What rules govern transfers of rights? Is there anything peculiar to STM works that warrants a call for no copyright protection or for different rules than those that govern other works of authorship? What are the key principles or values of the scientific enterprise that should guide the management of such rights?

(2) Economic and non-economic rewards to authors from STM publishing. What are the current and potential non/economic rewards to STM authors from publication, in both traditional and alternative venues? Can or should authors get financial compensation for their articles? Why has this been abhorrent in STM? How else, apart from monetary compensation, can authors share in the value of their work? The Social Science Research Network (SSRN), an online publisher of abstracts and (when authors can provide it) full text for working papers or forthcoming published articles, will be used as an example to explore these and other related issues in more detail.

(3) Licensing by authors to publishers, and from publishers to users. What rights can authors realistically give to or withhold from publishers? What are some alternative forms of managing rights, such as broad license to publishers while retaining copyright, and what are the pros and cons? What are major issues between the publishers and users of STM publications and of various models of access? Institutions generally gain access to electronic information via

license. What kinds of access do these licenses offer? How could access be improved? What are the advantages and disadvantages of licensed access models compared to hard copy distribution? What role might the science funding agencies or other organizations play in resolving conflicts in relationships between authors and publishers, or publishers and users?

Panel 4: What is Publishing in the Future?

The manufacturing and distribution machinery for digital documents (more generally *digital objects*) are readily available to many individuals and institutions through computer-based tools and the World Wide Web. The dominant remaining challenges and opportunities are the processes for evaluating and credentialing these digital objects, brokering interchange between the producers and consumers, and providing long-term access for those objects deemed worthy of such continuing investment.

Overarching this, especially in the science, technology and medical realm, is a blurring between the processes of creating data, information, and knowledge and the processes of publishing and use by others. More and more discovery and learning in the STM sector is being done in organizational forms based on cyberinfrastructure in which most of the raw ingredients, intermediate artifacts, and the processes of human interaction are in digital form. (These environments go by various names: collaboratory, e-science community, grid-community, etc.)

Scholarly communication in these communities is emerging in more of a continuous flow rather than the more discrete batch processes of the paper-and-ink publishing world. Publications evolve and can be shared at various stages of completeness and review. Some publications are artifacts representing the state of collaboration and are in a real sense never intended to be finished. The networks are becoming not just places where pre-credentialed documents are organized and made available, but rather technology mediated organizations – social structures – that are *credentialing.*

This panel will present and discuss some of these trends, new options, and implications including institutional repositories, pre-print servers, reputation and rating systems, open archives movements, and technology-based, self-organizing communities. The participants are Hal Abelson, Rick Luce, and Paul Resnick. Dan Atkins is the organizer and moderator.

Panel 5: What Constitutes a Publication in the Digital Environment?

The practices defining the collection, validation, representation and dissemination of scientific and scholarly knowledge have changed greatly in the past few years with the large scale introduction of curated community databases, the development of massive shared observational data archives, and the use of complex multi-component simulation models to conduct experiments in silico. We have seen growing focus on extremely complex multi-disciplinary research projects that seek to link together and integrate many different disciplinary information resources to understand phenomena as complex as global warming. Advanced information technology and networks allow scientific and scholarly communities to create, access and share such resources worldwide. These developments in turn are beginning to have profound effects on how we think about documenting and communicating scientific knowledge and research.

This panel complements Panel 4 by focusing on the changing nature of the information that is being created and disseminated, and how these changes alter the characteristics and conceptions of constructs such as the scholarly article and the scholarly journal. We will try to look at a range of perspectives. At the level of individual authors, we may be seeing new thinking about what constitutes a scientific paper and about how one's published "article" may relate to other materials such as datasets or computer programs. At the more systemic level, both articles and journals now live and develop within a ever more complex ecology of other information resources, and the linkages among the various components are becoming ever richer and more complex. This has implications for how editors and publishers as well as authors may need to think about the design and operation of journals in the future, particularly as electronic journals become

much more than simply images of printed pages that are stored and transmitted using information technology.

Clouds on the Horizon

Commercialization Constraints

The efforts of universities and faculty members to capture and exploit the soaring commercial value of the intellectual property created by research and instructional activities create many opportunities and challenges for higher education. There are many signs that the commercialization of intellectual property poses significant risks as well as benefits to research universities.

Today almost everything is viewed as having commercial value, be it a reagent, a research method, a clone of cells, a DNA molecule, or its sequence. Not only the results, but even the tools of science are now being restricted. In the absence of standard policies, industry can demand greater control over the research agenda, the release of research results, rewards to the institution and faculty, and the ownership of intellectual property, triggering competition among universities for corporate support of faculty research on the basis of customized conflict of interest agreements.

Today scientists sign agreements requiring them to keep both the methods and the results of their work secret for a certain period of time. More than a quarter of US geneticists say they cannot replicate published findings because other investigators will not give them relevant data or materials. There is growing evidence suggesting that industrial sponsorship actually influences the outcome of scientific work.[6] Universities are encountering an increasing number of conflict of interest cases, stimulated by the exploding commercial value of intellectual property and threatening not only institutional integrity but even human life in conflicted clinical trials.

On occasions some institutions and faculty members have set aside the most fundamental values of the university, such as openness, academic freedom, and a willingness to challenge the status quo, in

order to accommodate this growing commercial role of the research university.[7]

Donald Kennedy made an excellent further point in a recent editorial in *Science*. He suggests that just as the Vannevar Bush's *Endless Frontier* changed fundamental science from a venture dependent on small privileged elites into a vast publicly owned enterprise, Bayh-Dole and related federal policies is driving university research toward the private sector, fueled by the mobilization of philanthropy and corporate risk capital.

Continuing the frontier motif, he suggests we might regard the current framework characterizing technology transfer as the "Great Enclosure". Just as the Homestead Act of 1982 transformed the American frontier from public land into a checkerboard of individually owned holdings by allocating land virtually free to those who would promise to live on and improve it, the largely public domain of basic research is now moving into private hands by yet another federal act, Bayh-Dole, that allows universities or individual scientists to claim ownership of the intellectual property created by federally sponsored research. Interestingly, these enclosure revolutions came about in the same way: both were implemented by purposeful government intervention, accomplish through statute.

Kennedy contends that while this has brought some major benefits, it has also been accompanied by significant costs. New problems of conflict of interest, royalty distribution, and the propriety of commercial relationships have arisen for faculty members and university administrators alike. The contemporary enclosure of the Endless Frontier is replicating the history of the Homestead Act, yielding patent disputes, hostile encounters between public and private ventures, and faculty distress over corporate deals with their universities. Sometimes government action is unintended, such as the recent Executive Order on stem cell research that promises to transform a major public program into the propriety sector. Many observers, noting these costs, advocate policies for reversing privatization.

Homeland Security

After the September 11, 2001, assaults on the World Trade Center and the Pentagon, and the subsequent anthrax attacks via the postal system, the scientific, engineering, and health research community was quick to respond at many levels, from initiating new research to analyzing needs for improved security. This community recognizes that it has a clear responsibility to protect the United States, as it has in the past, by harnessing the best science and technology to help counter terrorism and other national security threats.

In meeting this responsibility, the scientific, engineering, and health research community also recognizes a need to achieve an appropriate balance between scientific openness and restrictions on public information. Restrictions are clearly needed to safeguard strategic secrets; but openness also is needed to accelerate the progress of technical knowledge and enhance the nation's understanding of potential threats.

A successful balance between these two needs -- security and openness -- demands clarity in the distinctions between classified and unclassified research. We believe it to be essential that these distinctions not include poorly defined categories of "sensitive but unclassified" information that do not provide precise guidance on what information should be restricted from public access. Experience shows that vague criteria of this kind generate deep uncertainties among both scientists and officials responsible for enforcing regulations. The inevitable effect is to stifle scientific creativity and to weaken national security.

To develop sharp criteria for determining when to classify and/or restrict public access to scientific information, as well as to address the other important issues outlined below, we call for a renewed dialogue among scientists, engineers, health researchers and policy-makers.

Achieving the purpose of scientific and technological activity -- to promote the welfare of society and to strengthen national security -- will require ingenuity from our science, engineering, and health community, as well as from the many agencies of the federal, state, and local governments involved in counterterrorism. The nation's safety and the continued improvement of our standard of living depend on careful, informed action on the part of both governments and the scientific, engineering, and health community. A continuing, meaningful dialogue needs to begin -- one that produces a true collaboration for the many decisions that need to be made.

A Final Observation

In sum, STM journals have put uncommon pressure on library budgets over the past ten years but the rate of increase has slowed

considerably in the past three years so there's some basis for being optimistic about cost management over the next ten years. {There is also an active dialogue among the publishers, vendors, aggregators, and libraries trying to further address the multiple pressures in the market place - publishers wanting to maintain their cash flow, subscription vendors working to define a new role in the era of publisher packaged e-journals, and libraries struggling with a major economic downturn seriously impacting their collections budgets in many institutions.

Let me leave you with a famous quote:

"If nature has made any one thing less susceptible than all others of exclusive property, it is the action of the thinking power called an idea, which an individual may exclusively possess as long as he keeps it to himself; but the moment it is divulged, it forces itself into the possession of everyone, and the receiver cannot dispossess himself of it. That ideas should freely spread from one to another over the globe, for the moral and mutual instruction of man, and improvement of his condition, seems to have been peculiarly and benevolently designed by nature, when she made them, like fire, expansible over all space, without lessening their density at any point, and like the air in which we breathe, move and have our physical being, incapable of confinement or exclusive appropriation. Inventions then cannot, in nature, be a subject of property." (Thomas Jefferson)

Not a bad principle to guide this two-day symposium!

---

1 Peter F. Drucker, "The Age of Social Transformation," *Atlantic Monthly*, November 1994, 53–80; Peter F. Drucker, *Post-capitalist Society* (New York: Harper Collins, 1993).

2 Erich Bloch, National Science Foundation, testimony to Congress, 1988.

19. Robert Lucky, "The Quickening of Science Communication", Science Vol. 289, July 14, 2000, pp. 259-264.

16. "Books, Bricks, and Bytes," *Daedelus* 125, no. 4, (1996), v-vii.

5 The co-principal investigators of the National Academies project are William A. Wulf, President of the National Academy of Engineering and Professor of Computer Science at the University of Virginia and James J. Duderstadt, Professor of Science and Engineering at the University of Michigan. Raymond E. Fornes, Associate Vice President for Research at North Carolina State University, has served as the senior Academy staff member for the project.

6. "Data Hording Blocks Progress in Genetics", Science, Vol 295, January 25, 2002, p. 599.

7. Eyal Press and Jennifer Washburn, "The Kept University", The Atlantic Monthly, March, 2000, pp. 39-54.