



The US government spends close to \$50 billion every year on scientific research (other governments and public institutions across the world more than double this figure), and hundreds of thousands of bright and dedicated scientists devote their careers to carrying out biological and medical research

The goal of this tremendous investment in basic research is to improve health, economic productivity, and the quality of life - both material and intellectual - of citizens of this country and the world

The product of this research - what this \$50 billion buys us - is a treasury of knowledge - new discoveries and new understanding of our bodies and the world around us, and new ways to improve our health and to prevent and treat diseases

The primary repository of this knowledge is  
the published, peer-reviewed scientific  
literature - the only permanent, public  
record of our ideas, results and conclusions  
and those of our colleagues and  
predecessors

# Impact of the Internet

The rise of the internet and the advent of electronic publishing have tremendous potential transform the way we communicate and use scientific knowledge - to make this fantastically rich but extremely fragmented and unsystematic information contained in the scientific literature far more accessible and useful

# Public Libraries of Science:

## GenBanks of the Scientific Literature

It is now possible to imagine the creation of electronic "Public Libraries of Science", open repositories containing the full-text of every scientific paper ever written in a common, structured format, freely available for access, downloading and use by anyone, anywhere

Important decisions are being made today that will determine whether such a vision is reality and will shape the future of our interaction with the scientific literature

It is absolutely critical that we - the scientific community - actively engage in this process to ensure that the future shape of scientific publishing best serves



# Overview of Talk

- Present a vision of a free and open scientific literature
- Discuss the ways in which this can improve the way we communicate and conduct scientific research
- Contrast with current practice
- Contrast with where scientific publishing is heading
- What we have done and what we can/should/must do

# Open Access Scientific Literature

Successful creation of public libraries of science requires that the full-text of published scientific manuscripts be, for all intensive purposes, in the public domain

This means that anyone can read, download, search, include in databases, redistribute or otherwise use the full-text subject only to the restriction that proper citations be maintained

# Benefits of Open Literature

The first, and most obvious, benefit of an open scientific literature will be unfettered access to the complete scientific literature for anyone, anywhere in the world

# Benefits of Open Literature

Currently, comprehensive access to the literature is available only to researchers at large, well-funded academic or research institutes in the developed world

# Benefits of Open Literature

Public libraries of science would enable researchers in poorer countries, or at smaller/poorer institutions in the developed world to have the same level of access to the scientific literature as a research at Stanford, Harvard or Cambridge

# Benefits of Open Literature

However, better access is only the beginning of the benefits an open scientific literature could provide to the scientific community and the world

# Benefits of Open Literature

Public libraries of science would enable creative researchers around the world to begin to tackle the challenge of building tools to better search and connect the treasury of information in the scientific literature, to link this information to other forms of knowledge (such as sequences), etc...

# Lessons from GenBank

The impact that GenBank has had on research involving sequences provides a useful example, as GenBank (and other similar databases) are in essence public libraries of sequence information



# Lessons from GenBank

The transformation of the life sciences by DNA sequences and the rise of genomics was absolutely dependent upon free and open access and unrestricted use of published DNA sequences - upon the ability to copy and use and transform and redistribute the information without any real restrictions imposed by producers or journals in which the sequences were published

# Lessons from GenBank

As sure as GenBank enabled and inspired the creation of methods and tools for sequence analysis upon which we and most of the scientific community are now dependent, robust public libraries of scientific knowledge would set off a boom in ideas and tools for accessing and using this information

# Lessons from GenBank

However, as sensible and natural as GenBank seems, it is critical to realize that the distribution of sequence information need not have followed this model

# Lessons from GenBank

Imagine how much of the scientific progress of the past decade would have been sacrificed if the publishers had treated DNA sequences as they do all other published information.

## Lessons from GenBank

Many publishers undoubtedly now wish they had claimed copyrights on the sequences they published. They would have been a lot richer now, but science (not to mention the biotechnology and pharmaceutical companies for whom the public sequences, and the tools and discoveries that have sprung from them, are the critical resource) would have been a lot poorer.

## Lessons from GenBank

Now consider the possibility that we may be sacrificing at least as much progress as was and is enabled by GenBank, by allowing publishers to prevent any similar creative use of all other published information - a much larger and richer body of information than the sequences in GenBank

# Will Public Libraries of Science become a reality?

The major obstacle to the creation of public libraries of science are scientific journals who persist in claiming ownership - through copyright and other means - of the scientific literature and who exercise this control to restrict access and use

# Will Public Libraries of Science become a reality?

The infrastructure for comprehensive archives of the scientific literature already exists and is constantly expanding and being improved, and virtually all scientific literature is produced in electronic form suitable for submission to archives (SGML/XML is heavily used in production)

The problem is that only a tiny, tiny fraction of this literature is going into the archives



# Ownership of Scientific Literature

*Certainly, by no reasonable standard can journals claim to have earned the right to exercise such ownership over the scientific literature*

*Their contribution, as important as it is, pales in comparison to that of the intellectual and physical input from people who did the work and the financial support of the public and private bodies that supported them*

# Ownership of Scientific Literature

The only question to ask is whether journal ownership and control is a necessary evil

# History of Scientific Publishing

# History of Scientific Publishing

Scientists have historically relied on paper publication as the most efficient and practical means for wide distribution and promotion of their work, and printed scientific periodicals have been the major means of carrying out this practice

# History of Scientific Publishing

Since the major costs in this system are printing and distribution, with each copy produced and distributed involving an expense for the publisher, a standard business model evolved in which scientific journals derived their income from selling periodical subscriptions to individuals and institutions interested in the topics covered by the journal

# History of Scientific Publishing

Since the production of a printed journal involved a significant investment on the part of a publisher, it was reasonable for them to seek to protect this investment by asking authors to assign them copyright on published works

# History of Scientific Publishing

This transfer of copyright to journals facilitated the publication process and came at a limited cost to individual scientists and the scientific community

# History of Scientific Publishing

The distribution of an author's work was not limited by copyright, but rather by the cost of printing and distributing copies and readers' or their institutions' ability to pay the legitimate cost of those copies



# History of Scientific Publishing

The ability to find information in the huge body of published scientific work, or to map and record connections between bits of information published in separate works, in separate journals, was also not limited by the business model or by copyright, but rather was inherently limited by the physical nature of the paper literature - serial publications in physically dispersed volumes.

# History of Scientific Publishing

Although this system was not perfectly fair - individuals and institutions who could not afford subscriptions were cut off from the latest scientific knowledge - given the inherent limitations of printed matter for distribution and organization of information, this system was arguably the most rational and efficient possible, and it served scientific authors, their readers and society well

# Scientific Publishing Today

In the digital age, none of the sound premises of this system remain valid, and the business model that served science so well in the era of printed journals has become a major impediment to progress.

# Scientific Publishing Today

Today, the costs involved in scientific publishing today are almost entirely in the preparation of the original edited electronic document - the original is as expensive to produce as ever, but the costs to produce and distribute each additional copy are now infinitesimal.

# Scientific Publishing Today

A business model that charges readers for each copy of a work is economically irrational and inefficient, and perversely thwarts that goals of authors, readers and the funders of the work by charging a high price for copies that cost nothing to produce or distribute, thereby artificially creating a barrier to the distribution of

information

# Scientific Publishing Today

Of course, the remaining costs of publishing

- organizing and maintaining an editorial board, managing peer-review, and turning submitted manuscripts into edited, marked up, formatted documents - still must be paid

# Scientific Publishing Today

However, since these costs all scale largely with the number of manuscripts submitted and published, it makes far more sense for the scientific community to pay these costs at the time of publication and to place the finished product in the public domain

# Scientific Publishing Today

The public and private agencies that support scientific research should view these costs for communicating the results of the research they funded as a final, indispensable part of the research process, to be paid upfront so that the knowledge produced by this research can be freely available to all



# Scientific Publishing Today

There is plenty of money to go around - these same institutions (governments, universities, foundations, companies) already fund the publishing process - they just do it indirectly through overhead or other money that goes to libraries to pay for subscriptions

# Scientific Publishing Today

Many billions of dollars are spent every year on scientific journals - more than \$10,000 per published article.

Doesn't it seem eminently reasonable to demand that in exchange for giving them the content for free, voluntarily providing most of the essential labor in this process and transferring an immense amount of money to the journals, that the finished product belong to the public rather than to them?

*In the end, no one can reasonably argue that science or the public interest is better served by limiting access to the information voluntarily published in scientific and scholarly journals, or restricting the ways it can be used*

# Who Benefits from an Open Scientific Literature

Everyone - scientists both as authors and readers/users of the literature, the institutions that fund their research, biotech and other research companies and the public would all benefit materially from a world where all scientific literature is in the public domain

# Who Benefits from an Open Scientific Literature

## Scientists as Readers/Users

The most immediate beneficiary of an open scientific literature will be scientists in our role as readers and users of the literature.

Think of all the things you could do if you could access, download and use the full text of every scientific article in a structured format.

# Who Benefits from an Open Scientific Literature

## Scientists as Authors

We are already giving away this work for free, and all we want in return is for our colleagues to know about our work and for us to receive proper credit (or glory!) for what we have done.

Clearly, removing barriers to access and use of our work can only help increase its dissemination and the likelihood that someone will take an interest in, and be able to read, what we have written

# Who Benefits from an Open Scientific Literature Funding Agencies/Institutions

Public and private agencies that fund research are interested in creating knowledge, and disseminating this knowledge to anyone who will be able to use it - for this purpose, the production of open literature clearly is preferable to a closed literature

Payment of the complete costs of publishing at the time and point of publication is also economically far more efficient and will, in the long run, save money.

# Who Benefits from an Open Scientific Literature

## Companies

Biotech companies do not have the institutional buying power to subscribe to many journals - an open literature would give them free access to any article they might need or want. They already paid for it once (though taxes); why should they have to pay for it again?

Will also create huge opportunities to develop commercial tools to help users navigate the open access literature.



# Who Benefits from an Open Scientific Literature

## The Public

Although most scientific research is paid for by the public

- through taxes - most people can not currently read articles describing research they paid for.

For example, today, a sick patient interested in reading about the latest research on their disease has to pay for access to each article. In most cases, this is also true for their family physician, most of whom subscribe to and have access to only one or two specialty journals.

# Who Does Not Benefit from an Open Scientific Literature?

## Existing Publishers

Many journals are immensely profitable, and even most non-profit publishers have become dependent on the revenues from these journals



## The Future: Costs of Complacency

Most scientists support the general idea of an open scientific literature, and see the likely benefits public libraries of the scientific literature could have on their work and on science, but there is a fairly common attitude of “well, the current system isn’t perfect, but it’s pretty good”.

## The Future: Costs of Complacency

To some extent, this is true. Electronic publishing as it exists today has made things much easier for many of us.

For example, from my computer in Berkeley, I can fairly rapidly find and access almost any article I want.

# The Future: Costs of Complacency

But I'm lucky. UC has subscriptions to virtually everything. Many people - including many of you probably - are not that lucky.

Furthermore the system is economically unstable. Journal costs are rising rapidly, and even wealthy universities are being forced to cancel subscriptions to many journals, imperiling our ability to comprehensively access the literature.

## The Future: Costs of Complacency

However, far darker scenarios are on the horizon. The trend amongst scientific publishers is towards content aggregation.

We have ScienceDirect and HighWire Press and others collecting huge amounts of the scientific literature in private archives.

# The Future: Costs of Complacency

This aggregation is a prelude to a major planned shift in how we will access - and pay for access - to these articles. The plans are all for a sophisticated system of authentication - where every scientist has a digital ID, and every time you read or access an article you will pay a small fee, which will, presumably, be paid by your



# The Future: Costs of Complacency

This might not sound so horrible at first glance (so long as you work for someone who can afford the charges) - no more passwords, and no more paying for articles that noone reads.

# The Future: Costs of Complacency

However, there are many reasons to fear this future. First, it will still be a virtual monopoly. If you want to read an article published in a HighWire journal, you will still have to access it through HighWire - you will have to pay whatever they want you to pay and you will be utterly dependent on the tools they provide to access the literature. It will be impossible for academic researchers to build tools to access and use the literature. There is no reason to suspect that Elsevier, Wiley, HW and others will provide the tools we want. For example, until as recently as a last year, people at HW were questioning whether scientists

## The Future: Costs of Complacency

The potential problems go further. Today, I can use Berkeley's electronic library to read essentially any scientific (or for that matter and academic) article. However, the move away from institutional subscriptions to accessibility based tolls will likely end this liberty.

# The Future: Costs of Complacency

Consider, for example, a researcher who studies mouse development who is funded by a grant from the NIH. They will likely put a line on the grant to cover the costs of their literature access. The NIH could quite easily – and in all likelihood would – approve, but would not be willing to pay for the scientist to read irrelevant articles.

This won't just restrict your ability to read art history, but quite possibly your ability to read articles in ecology, computer science, or other things that are not directly “relevant” to your grant, but which we all know are critical to performing the best science.

# The Future: Costs of Complacency

The key thing here is that private ownership of the literature + the developing internet architecture of authentication and control creates a highly regulable system, in which many of the essential characteristics of the scientific process may no longer be available

## The Future: Costs of Complacency

You may all thing I'm chicken little here, and screaming that the sky is falling, but this is the kind of thing about which publishers talk openly. Obviously, science is strong and this would not be the death of science, but think about all that we would lose if this became reality. Again, think about what sequence analysis would be like if similar rules applied

there

# The Future: Costs of Complacency

I hope that none of you want this world to become reality. However, there is often an immensely fatalistic tendency in how we think about journals.

We view journals like some mysterious, all-powerful, unalterable force, that decides how significant our work is, where we are going to get jobs and whether we will get tenure, and we are reluctant to mess with this force lest it punish us in unspeakable ways.

# We Control the Future

But we have to remember that we - the scientific community - hold all the power here. We do the research. We write the papers. We choose where to submit them.

We do the reviewing. We pay the costs. We are members of the scientific societies that publish many of the best journals. If we decided to change the system, we could do

it immediately



PubMedCentral

*"If you build it, they will come"*

PubMedCentral

*"It was built, and they (mostly)  
didn't come"*

Notable exceptions:

PNAS

BioMed Central

[Home](#)

[Open Letter](#)

[Read](#)

[Sign](#)

[See Signatures](#)

[What Next?](#)

[Journal Policies](#)

[PLOS Journals](#)

[PLOS License](#)

[PLOS Finances](#)

[Links to articles about PLOS, and related sites](#)

[Frequently Asked Questions](#)

The Public Library of Science is a non-profit organization committed to making the world's scientific and medical accessible to scientists and to the public around the world. We support scientific progress, education and the public good.

We are working for the establishment of international open access science that will archive and distribute the complete record of scientific articles, and foster the development of new ways to separate reports and segregated into thousands of different journals with its own restrictions on access.

As a step toward these goals, scientists around the world have circulating an open letter urging publishers to allow the complete record to be distributed freely to all public libraries of science. The response from the international scientific community to this initiative has been remarkable, and our open letter has now been signed by **29367** of you. Our initiative has prompted some significant changes in many scientific publishers towards free access to publication. In general these steps have fallen short of the reasonable level advocated. We will make every effort to publish our own journals with support to, those journals that have adopted the policy of open access.

# PLOS Open Letter

We support the establishment of an online public library that would provide the full contents of the published record of research and scholarly discourse in medicine and the life sciences in a freely accessible, fully searchable, interlinked form. Establishment of this public library would vastly increase the accessibility and utility of the scientific literature, enhance scientific productivity, and catalyze integration of the disparate communities of knowledge and ideas in biomedical sciences.

We recognize that the publishers of our scientific journals have a legitimate right to a fair financial return for their role in scientific communication. We believe, however, that the permanent, archival record of scientific research and ideas should neither be owned nor controlled by publishers, but should belong to the public, and should be freely available through an international online public library.

To encourage the publishers of our journals to support this endeavor, we pledge that, beginning in September, 2001, we will publish in, edit or review for, and personally subscribe to, only those scholarly and scientific journals that have agreed to grant unrestricted free distribution rights to any and all original research reports that they have published, through PubMed Central and similar online public resources, within 6 months of their initial publication date.

# Response to PLOS Open Letter

Signed by over 30,000 scientists

A few journals responded positively (NAR, Bioinformatics, ASM), but by and large, the response from established journals ranged from dismissive to hostile

# Response to PLOS Open Letter

Many journals have taken a tiny step, namely they make their back-content available for free access at their website

It is important to note that there is a fundamental and critical difference between this and open access

Consider if DNA sequences were only accessible at the website of the sequencing center that produced it, available for download one at a time, and searchable only through tools provided by the producer

# PLOS: Doing it ourselves

The warm response from the scientific community has convinced us that scientists want to publish in open-access journals

It is also clear that they largely believe (correctly or incorrectly) that their careers will be harmed by publishing their best work in extant open-access journals

# PLOS: Doing it ourselves

Since extant publishers have been largely unwilling to provide open access journals to the community, PLOS has decided that the only way for the scientific community to produce successful high quality, open access journals will be to create them ourselves



# PLOS Publications

Therefore, PLOS has decided to launch a publishing endeavor devoted to the open-access publication of scientific research according to a “pay upfront” business model.

Our primary goal is to provide a venue for people to publish their work and place it in the public domain.

Our secondary goal is to demonstrate the feasibility of our business model, with the hope that other journals will emulate our success and that we will reach comprehensive open access publishing as soon as possible.

# PLOS Publications

*We will begin with two journals devoted to works of “significance” - PLoS Biology and PLoS Medicine - to be followed by journals devoted to publishing any scientifically rigorous work.*

*We have been aggressively trying to raise money from foundations to launch these journals with sufficient quality, integrity and financial stability to ensure their success, and should be running by January 1, 2003.*

# The PLOS Model

*Publishing process like existing journals (i.e. works will be peer-reviewed)*

*All costs will be covered by up-front charges (\$500-\$1000 and decreasing over time)*

*Published works will be made available under PLOS license at the moment of publication as HTML, PDF and XML, places into any database that wants content and made available for individual or bulk download*

# PLOS License

Unlimited right to access, use, redistribute  
in whole or in part, subject only to the  
constraint that the original citation be  
maintained

## Related Activities

Raising money to obtain rights to and digitize all of the previously published scientific literature

Fostering efforts to develop tools to use the open access literature

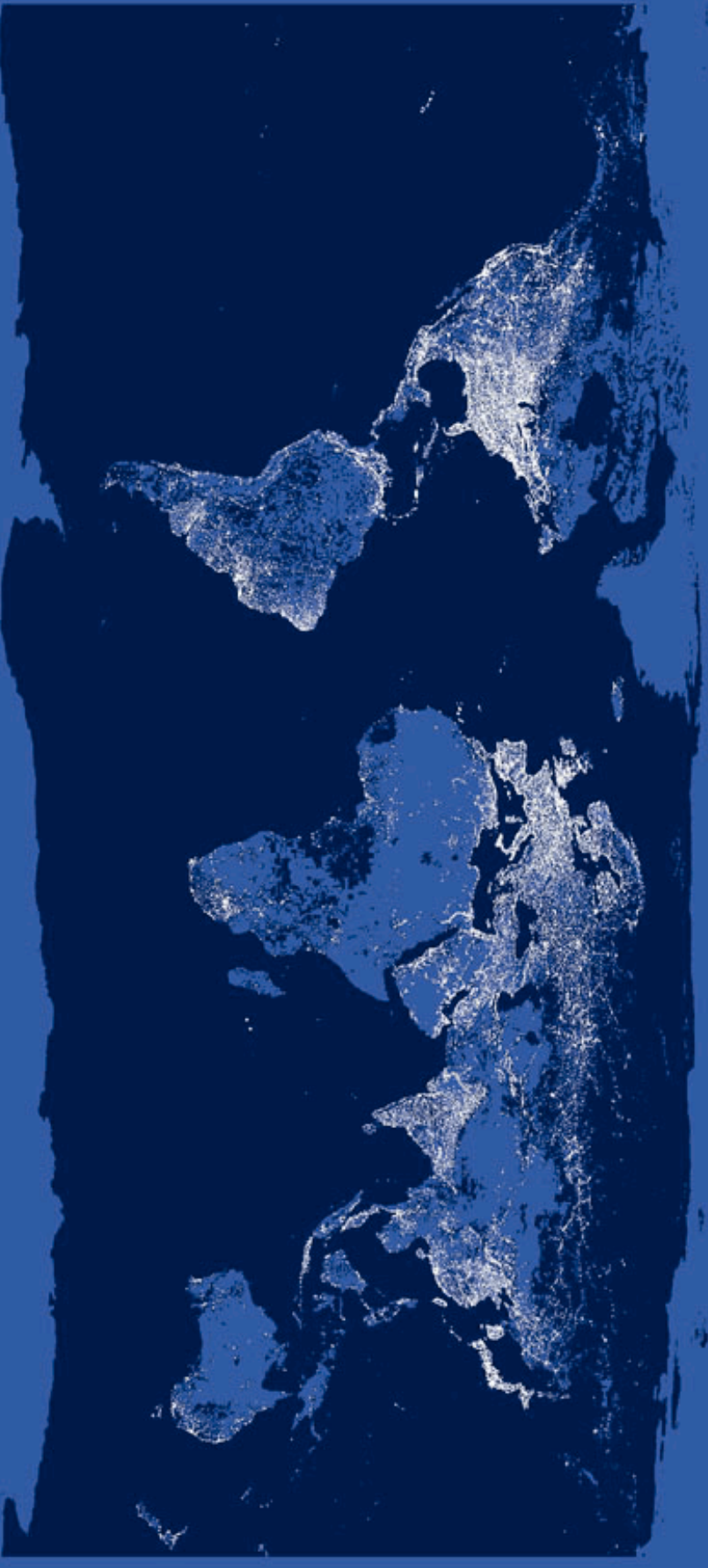
# We Need Your Help

Just as we believe that the scientific literature belongs to the scientific community and the world, these journals too will belong to the scientific community

To make them successful, we need your support and help. We need editors. We need reviewers. We need open-source software to manage peer-review and document production.

Most of all we need your best papers

# Oppose PRIVATE CONTROL OF Scientific Literature



*Public Library of Science*

[WWW.PUBLICLIBRARYOFSCIENCE.ORG](http://WWW.PUBLICLIBRARYOFSCIENCE.ORG)