

The 6th SPARC Japan Seminar 2009 (The 11th Library Fair & Forum 2009, Scientific Information Open Summit)

NIH Public Access Policy

This presentation will provide an overview of the US National Institutes of Health (NIH) Public Access Policy. The NIH is America's primary government agency for conducting and supporting medical research. It awards funds to scientific institutions to conduct biomedical research. The NIH Public Access Policy ensures that the public has access to the published results of NIH funded research. It requires scientists to submit final peer-reviewed journal manuscripts that arise from NIH funds to the digital archive PubMed Central (PMC) upon acceptance for publication. To help advance science and improve human health, the Policy requires that these papers are accessible to the public on PubMed Central no later than 12 months after publication. These papers are privately copyrighted, and are accessible under fair use. They cannot be downloaded in bulk, redistributed, turned into new works, etc.

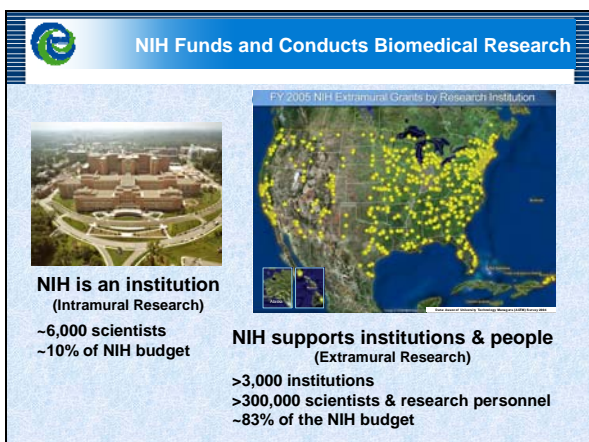
This policy was issued as a request to NIH funding recipients in May 2005. It became a requirement in April 2008. Making Public Access a requirement has approximately tripled collection of papers, from 19% of papers to almost 60% of papers. NIH continues to implement the Policy. Hundreds of scientific journals signed agreements with NIH to post all of their final published articles directly to PMC. NIH has also developed other ways in which authors and publishers may deposit author manuscripts to PMC.

Neil M. Thakur, Ph.D.

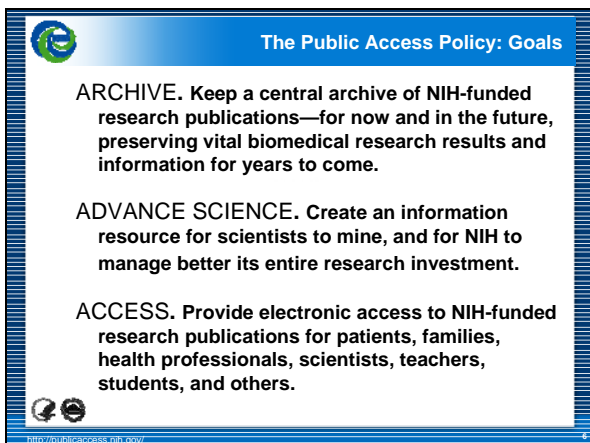
Special Assistant to the National Institutes of Health (NIH)
Deputy Director for Extramural Research

Introduction to the National Institutes of Health and Public Access

The National Institutes of Health (NIH) are the United States' primary funder of biomedical research with an annual budget of 30 billion dollars. The NIH employs 6,000 scientists on its campus and supports 3,000 institutions and universities mostly within the United States, but also some abroad.



The first thing I want to make clear is the difference between open access and the NIH public access policy. Open access refers to a movement to make scientific articles available for use in lots of different ways. These ways include a license, for example, to copy, to use, to redistribute or to reword and repurpose a scientific paper. The public access policy, on the other hand, is about access under fair use principles.. The work on PubMed Central is privately owned and we have permission to make that paper available without charge, but only for fair use. It is publicly available and it is privately owned. For example, that means that you can pass on a link to a paper on the NIH archive PudMed Central, but you could not make a copy of that paper and post it to your website. You would also need permission from the copyright owner to copy a figure or diagram and include it in your work



The Public Access Policy: Goals

ARCHIVE. Keep a central archive of NIH-funded research publications—for now and in the future, preserving vital biomedical research results and information for years to come.

ADVANCE SCIENCE. Create an information resource for scientists to mine, and for NIH to manage better its entire research investment.

ACCESS. Provide electronic access to NIH-funded research publications for patients, families, health professionals, scientists, teachers, students, and others.

<http://publicaccess.nih.gov>

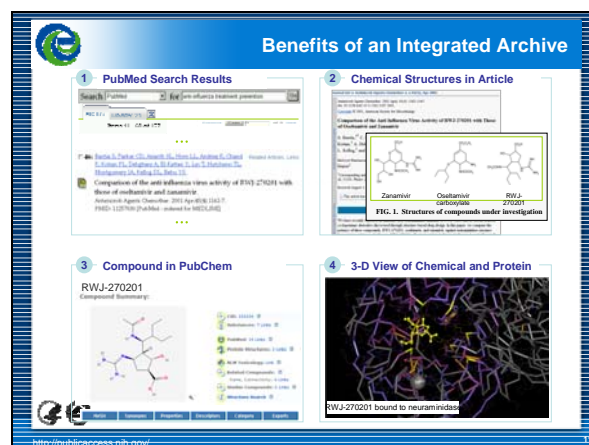
The NIH public access policy has three goals. The first is to create a permanent and stable archive for all the work that has been supported by NIH. The second is to advance science in two ways; by allowing all scientists to learn from NIH research results, and helping NIH better monitor and manage its portfolio of research investments. NIH supports about 86,000 papers a year and we do not have free access to these papers. For us to understand what it is that we are funding and how what we are funding is working, we need to have this archive and this policy. Also, since we have funded this research, we want all scientists to be able to use the results of our research to conduct their own work. Finally, all end users of this research should have access to this work as well. It is not just scientists and universities, but patients, doctors, engineers, students, and teachers.

This public access policy is derived from a 2007 law signed by Congress that requires the NIH to collect peer-reviewed manuscripts of funded research. It tells us that we are supposed to collect final peer-reviewed manuscripts, and sometimes this is referred to as the author manuscript. The law reads: “these manuscripts should be submitted on acceptance for publication”; that “these manuscripts should be made public no later than 12 months after publication”; and that “they should be made public on an NIH archive called PubMed Central.”

As a reference, PubMed Central is an archive of full text papers and PubMed is an index of abstracts used to find papers. PubMed Central has about 1.9 million papers and PubMed has about 18 million abstracts.

PubMed Central's Integrated Archive Approach

The NIH believes that the public access policy will help them advance science and improve public health. The American public expects that NIH-supported papers are available without charge to them. We believe this will advance science by posting papers to PubMed Central. PubMed Central is apart of a network of databases that are controlled by NIH. When a paper is on PubMed Central, it is integrated and accessible through these other scientific resources. We hope this additional access makes our research easier to read, easier to find, and easier to think about and build from. When NIH makes an investment in a particular area of science, it is easier for all scientists to work in that scientific area. NIH gets more competitive applications in that area, and we may see more publications and more research in that area. We expect more value for our money.



Benefits of an Integrated Archive

- PubMed Search Results**
Search for "anti-influenza treatment prevention" and find relevant articles.
- Chemical Structures in Article**
View chemical structures of compounds like Zanamivir and Oseltamivir within the article context.
- Compound in PubChem**
View detailed information for compound RWJ-270201 in the PubChem database.
- 3-D View of Chemical and Protein**
View a 3D molecular model of RWJ-270201 bound to neuraminidase.

<http://publicaccess.nih.gov>

Works hosted on PubMed Central contain links to the journal website, links to all the papers in PubMed Central that have cited this paper and links to key pieces of information that are in the abstract and the full text of the paper. You can go from a PubMed search for a paper that uses a particular chemical to look at the chemical structure in that paper, and then you can go to a chemical database, also run by the NIH, and in that database you will find related papers that also use that chemical. The NIH also has a 3D view of the chemical. The searches are not limited to words, but they also include concepts such as chemicals and genes and other things. We recalculate all of these linkages every

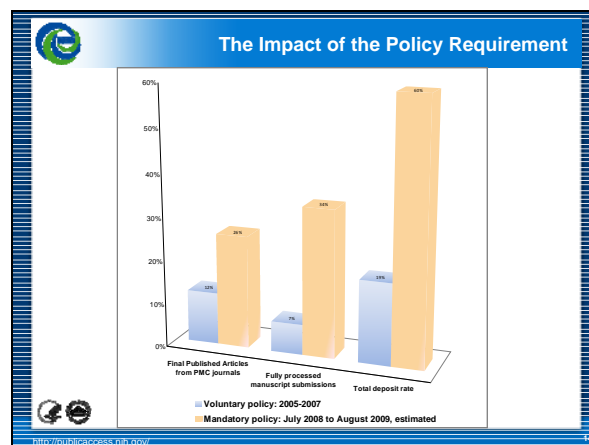
night. Every new paper makes our network both larger and more integrated.

Our policy specifically upholds copyright and asks all parties to uphold the principles of copyright. Authors can publish in any journal they wish and they can continue to transfer their copyright to publishers as they may have in the past. What we ask is that authors simply ensure that the paper is able to be posted on PubMed Central. They can transfer their copyright to the publisher on the condition that the publisher will ensure that the paper is posted, or the author can retain all of their copyright, or just a thread of their copyright to ensure that they have the ability to post their paper to PubMed Central.

The NIH Public Access Policy was voluntary from 2005 to 2007. It became mandatory through a law passed at the end of 2007 because the voluntary approach was not resulting in many papers being posted to PubMed Central.

The publishing community has amended their copyright policies and publishing agreements in response to our policy. Many publishers changed their copyright agreements to explicitly cover author manuscripts around 2005, 2004. Many American publishers or publishers that publish papers by American scientists now have specific provisions about the NIH public access policy.

In terms of our policy, NIH's formal discussions on public access policy started in 2004. We have had discussions about this topic and this issue for years before that. In 2004, we developed a draft voluntary policy and we held public meetings and we held a request for comments. After this comment period, we issued a voluntary policy in 2005, and we were only able to collect about 20% of our target number of papers. We eventually passed a law and made the public access policy a requirement. This requirement took effect in April 2008, and we are still implementing it.



Under the voluntary policy in blue, the NIH was able to collect almost 20% of its target number of papers. Under the current policy, it is now collecting about 60% of its papers.

Methods of Submission and Training Materials

There are two forms of papers and two ways in which they are being collected. One form is final published articles from journals. Publishers have signed agreements with us and they send us XML files. There is no author version of the paper posted to PubMed Central. Under the voluntary policy, we got about 12% of our papers through this method. We had about 300 journals partnered with us. Under the requirement, our numbers have jumped to 26% and currently include about 650 journals.

The second form is author manuscripts submitted through a web-based system. The author manuscript is taken in any format and converted into XML. The author confirms that the converted version is correct and the article is then made available on PubMed Central. The deposit of the author manuscript can be made by an author, someone from their staff, or by their publisher, but the process needs to be finished by the author. Only an author can confirm that the XML version of their manuscript to be made public is accurate and is their work. Under the voluntary policy, we got 7% of our papers this way, and now we get 34%.

The NIH makes training materials available in the form of PowerPoint slides available on their website. We ask institutions and universities to take these slides and give training support to their staff.

The first thing we ask authors to do is make sure they understand if their paper falls under the policy or

not. They need to make sure that the paper is peer-reviewed, it has been published or accepted for publication after April 2008, and is directly funded by NIH funding of some sort.

We also ask authors to think about and understand the copyright arrangement with their publisher. They need to make sure that any agreement they sign with their publisher allows them to have their paper posted to PubMed Central. They then have to make sure that their paper is deposited to PubMed Central.

Finally, they have to tell us that they are in compliance with the policy. When they send in their annual grant report, they must also include the research's PubMed Central ID number. When our NIH staff reviews these applications, we look for these IDs. When we do not see these IDs, we know that they are out of compliance, so we send them an email and we make sure that they get into compliance.

We find that authors write papers and they sign publication agreements, but they do not necessarily understand the details of the publication agreements. This spring, we started encouraging authors to think explicitly about six questions in their agreement as it related to the public access policy. How will the paper be submitted? What version of the paper will be made public, the author manuscript or the published article? Who will submit the paper; will it be the publisher or will

it be one of the authors? When will that paper be submitted, (because it is supposed to be submitted upon acceptance for publication)? What author is going to go onto the NIH website and approve that submission for posting? Finally, when will that paper be made public and what is the delay period? If an author understands these issues at the time of publication, then they will not have any problem complying with the policy.

After consultations with publishers, four methods of submitting papers to PubMed Central have been developed. These methods vary by what version of the paper is made public, whether it is an author manuscript of the final published article, whether the author has a role in the submission process, and whether the publisher sends the paper in directly and the author is not involved.

We have a website that provides all of this information. It includes training material for authors. It also includes policy information and background. We also run an e-mail help desk if authors have questions, but we do not have a telephone help desk. Our web-based submission system, the NIH Manuscript Submission System (NIHMS), has a separate website.

In this way, NIH is implementing its public access policy. Thank you very much.

Profile



Neil M. Thakur, Ph.D.

Since November 2005, Neil Thakur has served as a Special Assistant to the National Institutes of Health (NIH) Deputy Director for Extramural Research. He also serves as program manager for the NIH Public Access Policy. Prior to his time at NIH, he was Assistant Director of Health Services Research and Development at the Department of Veterans Affairs. Dr. Thakur holds a Ph.D. in Health Policy from Yale University School of Public Health and completed a NIMH postdoctoral fellowship in mental health services research at the Cecil G. Sheps Center for Health Services Research at the University of North Carolina at Chapel Hill.