9-8-2018

Open Access at MIT and Beyond: A White Paper of the MIT Ad Hoc Task Force on Open Access to MIT's Research

Katharine Dunn
Massachusetts Institute of Technology, khdunn@mit.edu

Hal Abelson
Massachusetts Institute of Technology, hal@MIT.edu

Chris Bourg
Massachusetts Institute of Technology, cbourg@mit.edu

Ellen Finnie
Massachusetts Institute of Technology, efinnie@mit.edu

Follow this and additional works at: http://digitalcommons.unl.edu/scholcom

Part of the Intellectual Property Law Commons, Scholarly Communication Commons, and the Scholarly Publishing Commons

Open Access at MIT and Beyond

A White Paper of the MIT Ad Hoc Task Force on Open Access to MIT's Research

MIT Massachusetts Institute of Technology

CONTENTS

ACKNOWLEDGMENTS

INTRODUCTION

MIT researchers, students, and staff have long valued and put into action MIT's mission to generate and disseminate knowledge by openly and freely sharing research and educational materials. Indeed, the Institute has been at the forefront of the sharing culture: MIT launched OpenCourseWare (OCW), a free web-based publication of virtually all MIT course content in 2001; in 2002 released DSpace, an open-source platform for managing research materials and publications co-created by MIT Libraries staff; and adopted the first campus-wide faculty open access (OA) policy in the US in 2009.

Convening an open access task force was one of the 10 recommendations presented in the 2016 preliminary report of the Future of Libraries Task Force. In July 2017, Provost Martin Schmidt appointed the MIT Ad Hoc Task Force on Open Access to MIT's Research (OA task force) to recommend ways that MIT's OA policies can be revised and updated "to support MIT's mission to disseminate the fruits of its research and scholarship as widely as possible." The OA task force is co-chaired by Class of 1922 Professor of Electrical Engineering and Computer Science Hal Abelson and Director of Libraries Chris Bourg, and includes a diverse and multidisciplinary group of faculty, staff, postdocs, and graduate and undergraduate students (see Appendix A).

The term "open access" as used by the task force comes from the 2002 Budapest Open Access Initiative and refers to research literature (typically journal articles) that is immediately, freely available on the public internet: Anyone would be able to "read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself."[1]

That said, open access is applied in different ways; i.e., OA doesn't always imply that you can reuse research articles for "any lawful purpose." Articles may be cost-free to read but still subject to publisher policies that limit other uses. Educational materials, data, and code, which MIT researchers routinely create and release publicly, may also be made openly available under different terms.

Open access may have begun simply because the web allowed for easy sharing, but it has evolved into a complex movement with political, social, and economic dimensions. The scholarly journal publishing system is unique in that researchers contribute their articles with no expectation of payment; at the same time, some publishers charge ever-increasing subscription fees, restrict authors' rights to reuse work, or both. Variations in the type of open access that will help "fix" the system are at the heart of debates among researchers, funders, librarians, and publishers.

This white paper is the first deliverable of the OA task force. Its goal is to give MIT students, staff, and faculty an overview of the open access landscape at MIT, in the United States, and in Europe to help inform discussions at the Institute over the next year. These discussions, which will take place at community forums and in other venues, including the task force idea bank, will help inform the task force as it develops a set of recommendations across a broad spectrum of scholarly outputs, including articles and books, data, educational materials, and code.

Part I of this paper provides an overview of current OA policies and movements in Europe and the United States as a way to give broader context to what open access means in practice internationally. Part II explores MIT researchers' approaches in terms of making their publications, data, code, and educational materials openly available.

---

[1] From Director of the Harvard Office for Scholarly Communication Peter Suber's blog post describing the Budapest Open Access Initiative definition of OA.

PART I: THE OPEN ACCESS LANDSCAPE IN EUROPE
AND THE UNITED STATES

MIT's open access practices exist within a complex and rapidly evolving landscape. This section looks at the ways in which different funding models, political structures, and priorities are shaping how open access is achieved.

Open Access in Europe

Open access was in many ways born and reared in Europe. The Budapest Open Access Initiative released its declaration in early 2002, defining two types of OA: self-archiving and OA journals.[2] Europe is also home to the Directory of Open Access Journals (DOAJ), SHERPA RoMEO (a database of publisher polices on article sharing), the Registry of Open Access Repository Mandates and Policies (ROARMAP), the Confederation of Open Access Repositories (COAR), and several of the main-player publishers, including Elsevier, SpringerNature, and Taylor & Francis.

Funders: United Kingdom

UK research funders have had OA policies for more than a dozen years, putting the UK ahead of most other countries. A few of the most important funders with OA requirements are the Research Councils UK (RCUK) and the Higher Education Funding Council (HEFCE), which both disburse public funds, and the private charity Wellcome Trust, which funds biomedical research. (In April 2018, RCUK and HEFCE merged to form UK Research & Innovation. Because the majority of policies discussed here originated with the former entities, their previous names will be used when referring to those policies.)

RCUK, an umbrella organization that gives public money to the country's seven research funders, has had open access requirements since 2005, the same year that Wellcome Trust launched OA policies. In 2012, an independent group commissioned by the UK government published a paper, known as the Finch Report, on how best to expand these and other OA requirements. The UK government accepted the Finch Report soon after it came out, and RCUK incorporated the report's recommendations into its revised OA policy in 2013.

The key change introduced by the Finch Report was a shift in focus from "green" OA (whereby an author "self-archives" manuscript versions to repositories)—similar to OA policies common in the United States—to "gold" OA (whereby publishers are paid to make the version of record open access). RCUK "prefers" gold OA via journals that are either fully open access or hybrids—paid subscription journals that offer a paid OA option for individual articles—and offers block grants to universities and research organizations to cover the cost of article processing charges (APCs) to make articles open. This means millions of pounds of public money that RCUK receives from the government goes to universities to pay for gold OA.

Gold OA papers funded by RCUK must be made immediately available under Creative Commons attribution (CC BY) licenses, the most liberal of the CC licenses, allowing even commercial reuse. (There is a range of CC licenses, which give authors a way to open their work to the public and allow different uses.) In the first academic year of the new RCUK policy, 2013–2014, block grants paid to make about 45% of RCUK-funded research papers OA; the goal for 2017–2018 was to raise that to 75%.

For data, RCUK states that "publicly-funded research data should be openly available to the maximum extent possible," and offers guidance on best practices in data management. The RCUK OA policy, however, has no data requirement.

---

[2] Subsequent statements further expanded on author rights (Berlin Declaration, 2003) and OA publishing (Bethesda Statement, 2003).

An overarching goal of the Finch Report was to push the global scholarly publishing system away from the model of library subscription payments and towards "OA publishing worldwide." The report's authors estimated that increasing access in this way—while still relying on subscriptions and repositories, at least for some time—would cost an additional £50–60 million per year in a system that spends £5.5 billion a year on research. To some, it is not clear that the investment is paying off. "We are simply not seeing that transition to OA amongst the traditional publishers that Finch envisaged," wrote Chris Banks, director of library services at Imperial College London, in 2016.

One factor potentially inhibiting a transition to worldwide open access is the way European funding is used. Though intended to advance open access, it has often been spent in support of hybrid OA models, where individual articles but not entire journals are made openly available. There can be a lack of transparency in how these additional open access fees are accounted for in relation to subscription costs. In March 2018, Wellcome Trust, which gave £822 million and £1.1 billion to researchers in 2015–2016 and 2016–2017, respectively, announced that it is reviewing its OA policy, in part because such a large proportion of its spending (71% in 2015–2016) goes towards OA articles in hybrid journals. Wellcome found that APCs for hybrid journals are 34% higher than those for fully OA journals. Hybrids have also been a concern because publishers are not always clear on how they're compensating for the fact that an article is, in effect, paid for twice (so called "double dipping")—through both the APC and a library subscription cost. Some publishers have reduced their subscription costs in response. UK Research & Innovation is also now reviewing its OA policies.

UK funders have pushed OA policies further still: In 2014, HEFCE began tying assessments of research quality—and its funding of university departments—to open access. In order to qualify to submit articles to the 2021 Research Excellence Framework (REF), journal articles and conference papers accepted since April 2016 must be deposited in an OA institutional or subject repository within three months of acceptance for publication. Results of the REF, in which institutions are scored on research output and impact, among other things, are used to determine how much money they get from the government. Therefore, there are direct consequences if authors do not make their work OA.

Funders: Europe

Major funders and universities located in other European countries—including France, the Netherlands, Sweden, and Germany—have their own OA policies or mandates. These were strengthened when, the day after the Finch Report came out in 2012, the European Commission (EC) announced Horizon 2020 (H2020), an €80 billion funding program with OA policies requiring that grant recipients make their articles OA via green or gold routes.

For green (self-archiving) manuscripts, the EC expects authors to figure out a way to comply.[3] "As the author, you are the legal copyright holder and can decide what to do with your copyright. This gives you the opportunity to request [sic] the publisher to reconsider his refusal given the EC Open Access policy, or submit your article to another (Open Access) journal, that enables you to fulfill your requirement," notes Open Access Infrastructure for Research in Europe (OpenAIRE), which will monitor research outputs from H2020. The EC created a model amendment to publishing agreements to help authors hold onto rights in order to meet the requirements. For gold OA, the EC will reimburse APCs between 2014 and 2020, the timeframe of H2020.

H2020 requires researchers to make data "as open as possible, as closed as necessary." This means that, though the default should be to make data openly available in a repository and under a Creative Commons license, the EC allows researchers to opt out of the policy.

In September 2018, 11 European funding bodies announced a bold initiative called Plan S, which will require researchers they fund to publish their articles in fully open access (non-hybrid) journals or OA platforms, starting in 2020.

---

[3] OA policies like MIT's, which use the legal mechanism of a license to let authors maintain rights to their papers, have not yet caught on in the rest of Europe. More details about UK efforts below, under "Proposed Policies and Movements."

Publishers and Open Access

European countries have led the world in terms of negotiating with (and standing up to) publishers. A survey published in April 2018 by the Brussels-based European University Association reported that 11% of negotiating consortia in Europe made deals last year that took into account open access publishing costs; 63% planned to do so in the future. Consortia of libraries, universities, and research institutes in Germany, the Netherlands, Finland, Austria, France, Sweden, and the UK have pushed publishers to varying degrees for license agreements that make their researchers' papers more openly available (sometimes asking for immediate OA, with no embargo) and offer sustainable price models. For the most part, the publishers involved in these negotiations are Elsevier and Springer, two of the three companies responsible for nearly half of published science papers, according to a 2015 study.

Germany has been in publisher negotiations around OA longer than most other countries. There are 200 institutional members of the Projekt DEAL consortium, which has asked publishers including SpringerNature, Wiley, and Elsevier to accept annual lump sum payments[4] that cover open access to papers whose corresponding authors are in Germany and online access to all publisher content for paying consortium members. The consortium also wants "fair" processing charges for OA articles. Projekt DEAL's goals are akin to the Finch Report's: to "trigger a big flip" to OA. "If it works, it would be a model for the rest of the world," negotiator Günter Ziegler of the Free University of Berlin told *Science* in August 2017.

The economics here are key, observed Bodo Stern, chief development and strategy officer at the US-based Howard Hughes Medical Institute, in a May 2018 issue of *EMBO Reports*. "We think the DEAL project is an essential initiative to realign payment streams with publishers. Subscription and open access fees are typically paid from different sources. Libraries pay institutional subscription fees, while funders and institutional research budgets typically pay open access fees. This separation of funding sources has allowed commercial publishers to tap into open access funds without reducing their subscription income. In principle, open access should reduce the overall cost of publishing, but this cost saving will not be realized until open access payments are offset from subscription payments and open access publishing grows at the expense of, not in addition to, subscription journals."

SpringerNature and Wiley are apparently open to the Projekt DEAL proposal and are currently negotiating licenses with the consortium. With Elsevier, however, negotiations have stalled. The publisher has not agreed to the "publish and read" model of combining OA publishing and access in one fee.[5] In July 2018, Projekt DEAL announced it had suspended talks with Elsevier; Elsevier cut the consortium's institutions off from access to its journals.[6]

---

[4] Projekt DEAL estimates a payment based on multiplying the number of papers with German first-authors by a "reasonable" APC per paper, which they consider to be €1,300–€2,000 per article. A Max Planck Digital Library policy paper estimates that subscription fees paid by the world's academic libraries work out to €3,800–€5,000 per paper.

[5] This is an example of "offsetting": As the total cost of publication grows for universities paying both subscription fees and APCs, offsetting methods and models have emerged as a way to remove duplicated effort. An early example of offsetting was the Royal Society of Chemistry's (RSC) "Gold for Gold" program, launched in 2013, in which libraries subscribing to the entire package of RSC journals received vouchers for their institution's authors to use to make their RSC articles open access. RSC now offers "read and publish" agreements that include making articles authored by a given university open access, along with providing subscribed access to paywalled articles. MIT Libraries recently signed a read and publish deal with the RSC.

[6] This is one of several stalled agreements: Consortia in Sweden, the Netherlands, and France have also failed to make deals with Elsevier, the Royal Society of Chemistry, and SpringerNature, respectively.

Proposed Policies and Movements

OA2020

In 2015, the Max Planck Digital Library in Munich published a paper outlining a proposal to achieve the "big flip" to open access that the Finch Report, German libraries, and others have been aiming for. OA2020, as the Max Planck initiative is known, asserts that the transition from subscriptions to OA "must start with libraries."

OA2020 supporters believe sufficient funds to make the transition are already there; money just needs to be redirected to OA publishing but in ways that reach beyond hybrid journals, offsetting, and APCs. OA2020 is asking libraries to commit to reallocating their full journal subscription budgets to "support sustainable, transparent, open access business models"—that is, to convert all subscription journals to OA and do so relying on many different economic models. More than 100 institutions have signed the OA2020 expression of interest, including 10 US universities.

United Kingdom Scholarly Communications License

Gold OA gets much of the attention in Europe, in large part because government funds pay to make journal papers openly available this way, and because of movements like OA2020. But, as noted above, European funders and universities also have policies for green self-archiving manuscripts. These policies are generally statements of support rather than MIT-style OA licenses, because authors remain beholden to publisher agreements.[7] Though funders like the European Commission offer contract amendments to researchers, most put the onus entirely on authors to hold onto rights. British librarians Chris Banks and Torsten Reimer have proposed a solution: the UK Scholarly Communications License (UK-SCL).

The license is much like MIT's and Harvard's (and, indeed, the authors cite Harvard's OA policy as inspiration[8]): Authors grant to their institution a non-exclusive license to the rights in their scholarly articles, and the institution makes the work OA under noncommercial Creative Commons licenses. The UK-SCL would restrict waivers to the policy for up to two years so that papers are eligible for the Research Excellence Framework.

Banks and Reimer argue the UK-SCL fulfills both funder requirements and the needs of "scholarly communications" writ large—that is, to share work—without compromising esteem because authors can still publish where they want. And, they say, it will cost less for universities than the gold model: "Hybrid publishing is not contributing to the transition to OA as envisaged by the UK Finch group, but is just eating up money—80% of our spend at Imperial," wrote Banks in 2016.

The model policy is still in draft stage, but some publishers have already pushed back. The London-based Publishers Association (PA) responded to the proposed license by saying that, among other things, it would cut into their earnings, impose a "very significant" administrative burden on researchers, and that "it conflicts with UK policy on open access, putting green open access before gold." "Publishers," the PA writes, "would have no choice but to require authors to obtain waivers."

---

[7] The MIT faculty open access policy operates through a license given to MIT to exercise all rights under copyright for a given article, except the right to sell the article for a profit. Because this license precedes any publication contract, it takes precedence over the publisher's standard policies.

[8] Reimer and Banks note that the UK-SCL is different from Harvard's in that it is adapted to UK law. In 2016, Banks wrote that during license discussions with funders the "moving wall" concept arose, "whereby, on deposit in a repository, metadata is made immediately visible and, over a period of time, the output itself transitions until it is available under a CC BY license. Depositing under the terms of this license would mean that outputs would be compliant with RCUK, HEFCE, and Wellcome requirements, and with the major European funders with OA mandates. At the same time, the citation advantage of early scholarly communication of the research findings would begin to accrue." Reimer does not mention this kind of gradual OA in his 2017 paper.

Open Access in the United States

Funders

In 2013, the Obama administration issued a directive requiring large federal agencies (those with more than $100 million in research and development expenditures) and their researchers to make articles and data from grant-funded research openly available within a year of publication. More than a dozen federal funders have released public access plans, and compliance is now just getting underway. (The exception is the National Institutes of Health (NIH), which has had sharing requirements since 2007 and enforces compliance.) All funders except the Department of Energy (DOE) require that authors submit peer-reviewed manuscripts of their papers to a designated funder repository. The DOE allows authors to use their home institutional repository; as a result, the MIT Libraries made a form to support authors depositing papers to DSpace@MIT under the DOE policy. MIT Libraries staff have been in regular contact with staff in the DOE's Office of Science and Technical Information since fall 2016 about best practices in terms of MIT author compliance.[9]

Many private funders also have OA policies, and these have in some cases directly affected publisher behavior. As of January 2017, the Bill & Melinda Gates Foundation requires its funded researchers to make papers immediately available under a CC BY license in the final published version. When Gates announced its policy, journals including *Nature*, *Science*, the *New England Journal of Medicine* (*NEJM*), and the *Proceedings of the National Academy of Sciences* (*PNAS*) did not offer this kind of publishing: they required embargoes, only allowed authors to post an accepted manuscript, or did not have CC BY licenses.

This misalignment between Gates' OA requirements and publisher options made it initially impossible for Gates-funded MIT authors to publish in these key journals. MIT's Office of Sponsored Programs and the MIT Libraries engaged in conversations and problem solving with Gates and the publishers about how to resolve this issue. In 2017 the American Association for the Advancement of Science (AAAS) and Gates ultimately reached a pilot agreement in which the foundation paid AAAS upwards of $100,000 so that authors could publish in *Science* under the terms of the Gates policy.

The pilot, which funded about 30 *Science* papers (from all institutions) and ended in June 2018, was not without its detractors, given that Gates offered substantial additional funding to achieve open access in this single prestigious journal—a model that raised questions about sustainability. "I'm glad to see it come to an end," said Peter Suber, director of the Harvard Office for Scholarly Communication. Suber called it a "prestige tax" that was "unnecessary and undesirable."

NEJM and *PNAS* have since changed their policies to support Gates' OA requirements. And, while *Nature* doesn't have a policy specifically for Gates-funded authors, its journals occasionally publish articles under CC licenses, including several Gates-funded ones.

There is much interest in what the Gates policy means more broadly: Funders are in a position to use their leverage to require publishers to accept terms to make authors' work more openly accessible. Perhaps in recognition of their influence, a group of a dozen private funders are members of the Open Research Funders Group, whose goals include developing ways to monitor compliance with OA policies and figuring out workflows that "minimize hassles" for grant recipients and administrators.

---

[9] Some issues have arisen with the DOE policy. For example, like all US funders, the DOE asks that the author's manuscript be made publicly available. However, the MIT Libraries has negotiated agreements with publishers, including the American Physical Society (APS, whose authors are heavily DOE-funded), in which APS sends the final version of record directly to DSpace@MIT. It would be burdensome to ask authors to track down a different version of a paper we already have. In September 2017, DOE agreed that final published version complies with its policy, as long as a given publisher is "friendly" in terms of sharing this version. (APS is friendly in this way.) Another example: in its publishing agreement Elsevier requires that authors opt out of open access policies, and it also has embargo periods on article posting that range from 12 months to four years. Authors can rely on MIT's OA policies to hold onto sufficient rights to meet DOE's requirements; however, if they opt out of the OA policy, and if there is an embargo longer than 12 months, authors cannot meet DOE's requirements. This issue remains unresolved.

Universities, OA2020, and the Article Processing Charge Debate

US universities have largely focused on green self-archiving rather than gold OA publishing,[10] but this may be changing. About 10 US universities, largely members of the University of California (UC) system, have signed on to the gold OA-focused OA2020 initiative.

When the initial OA2020 paper was released, librarians at UC Davis and the California Digital Library were two years into a project called Pay It Forward, which attempted to address the question of whether there is already enough money in the scholarly journal ecosystem to pay for a transition to complete open access. The report's authors reached three main conclusions:

- If North American research universities switch from paying subscriptions to paying APCs for all journals, they will pay more than they currently do.

- That said, the most research-intensive institutions receive grant money, and many researchers are already using these grant funds to pay publishing fees.

- APCs should decrease in the future, rather than increase, "as authors start to make decisions about where to publish based on a personal cost/benefit analysis."[11]

The study concludes that the question of how to set up a system where authors have "skin in the game"—and thus place downward pressure on prices—is key in terms of moving to a sustainable model for scholarly journal publishing.

University of California took another bold step following the Pay It Forward study, releasing a "call to action" in June 2018. In "Championing Change in Journal Negotiations," UC lays out strategies and principles for negotiating journal contracts with the goal to "end the system of journal subscriptions" and achieve an open scholarly communications system. Some of UC's guiding principles, drafted by faculty, include:

- No copyright transfers

- No double payments (one of the concerns with hybrid journals)

- No new paywalls

With UC laying out this strategic course to "pursue open access to the journal literature and a broader transformation in scholarly communication," there are still criticisms and questions about APCs and the emphasis on gold OA (which may explain why few US institutions have signed onto OA2020). These include worry that commercial publishers could consolidate their dominance in the market and charge what they want for APCs,[12] and concern that researchers from Global South or other countries with few funding resources may not be able to pay for APCs, may not qualify for waivers, or may lack centralized purchasing consortia to effectively negotiate deals with publishers to make their researchers' articles open access.[13]

---

[10] According to ROARMAP, there are about 70 universities, campuses, or schools with open access policies that focus on self-archiving the author's final manuscript.

[11] From a 2017 interview with Mackenzie Smith, University Librarian at the University of California, Davis, and principal investigator on the Pay it Forward team.

[12] As historian of science Jean-Claude Guédon writes, "Between 2006 and 2012, large commercial publishers had indeed come to the conclusion that they could safely embrace Open Access, at least a certain version of it: with OA, not only could they maintain their position in the publishing world, but they could even improve it by adding new business models to their old, print-inspired, subscription revenue stream."

[13] From Heather Joseph, executive director of the Scholarly Publishing and Academic Resources Coalition.

There are, certainly, OA strategies besides APCs and the "big flip." One example comes from funders. For example, the Gates Foundation and other funders have launched OA platforms[14] where their funded researchers can publish their work. Gates Open Research offers "immediate" publication and open peer review of articles.

There are also efforts to open access to individual fields, rather than (or prior to) attempting to convert the entire corpus of scholarly work. This is what happened when 3,000 libraries, funding agencies, and research institutions around the world partnered with publishers to convert high-energy physics journals to OA. The Sponsoring Consortium for Open Access Publishing in Particle Physics (SCOAP3) [15] gathers contributions from countries and centrally pays publishers to make journals OA under CC BY licenses. Discussions are now underway about how to transition fields such as linguistics and anthropology to OA, coordinated in part by the Association of Research Libraries (ARL).

Another strategy, on an even larger scale, is that libraries change their focus, reallocating acquisition funds towards the creation of a network of libraries in which they gather, store, curate, and preserve works created by researchers at their home institutions, and that these repositories link out to those at other institutions. For example, ARL's "academy-owned publishing" project promotes the idea that the "pathway toward the goal of open academic scholarship lies in developing platforms for publishing open, peer-reviewed journals that are managed by the academy in conjunction with scholarly societies and communities." The strategy herein is to offer authors alternative options to traditional journals (which are "too expensive, too slow, too restrictive, and dominated by entities using lock-in business practices") by developing publishing layers on top of preprint servers (which host pre-peer-reviewed papers) and other repositories. Overlay journals, which link out to articles in a growing array of subject-specific repositories, are an example of this strategy.

Related to this is the Confederation of Open Access Repositories' (COAR) next-generation repository work, which also calls for layering "value added services" (such as peer review) on top of repositories, and specifies technical requirements needed for repositories worldwide to be interoperable.

Open Access Trends and Tools

Alongside the large-scale initiatives and debates discussed above, there are smaller—but nonetheless important—trends and tools that have emerged in this space.

One of these is the rise of preprint servers or so-called "arXiv clones." In the last several years, more than a dozen new repositories of this type have emerged, some in fields like sociology (SocArXiv), biology (bioRxiv), and chemistry (ChemRxiv), whose researchers are still new to sharing pre-peer-reviewed work in this way. Most preprint servers and their users have in common the view that this kind of OA sharing will stimulate discussions, improve research, and open up science.

Researchers are also sharing work in different kinds of repositories—commercial ones like ResearchGate and Academia.edu—and the versions they share are not always legal, in that authors may not have the rights to

---

[14] A recent article in *EMBO Reports* argues that such initiatives make sense in a world in which companies like Elsevier are expanding their reach. Elsevier, for example, now calls itself a "global information analytics business" rather than a publisher. As Howard Hughes Medical Institute's Bodo Stern said: "I have been observing for many years that some of the large publishers have been internally setting a course towards a future when access to research articles is free. They are developing tools and systems to support the research workflow. This comes with many risks, for example lock-in effects, as well as tracking of users. So, in my view, academia should not repeat the mistakes made in the domain of scientific publishing many years ago. We should not outsource research assessment and evaluation to third party actors, like publishers."

[15] SCOAP3 is a worldwide consortium of libraries, funding agencies, and research institutions partnering with publishers to convert high-energy physics journals to OA under CC BY licenses.

put papers online. In September 2017, a group of more than 140 publishers wrote to ResearchGate asking the site to implement a system to help users determine the legal status of articles they want to share. A subset of those publishers, including Elsevier, Wiley, and the American Chemical Society, formed a group called the Coalition for Responsible Sharing and sent takedown notices regarding some of the 7 million articles that ResearchGate "illicitly" made freely available. After Elsevier and ACS filed a lawsuit in Germany against ResearchGate (which is based in Germany), the site removed some articles from public view.

A number of groups have built tools to help users unearth legal OA versions of articles when searching online. Unpaywall and OA Button are two of these: The tools can be added as browser extensions, and they search thousands of OA repositories and publisher sites to find OA versions of papers.[16]

---

[16] The MIT Libraries has added oaDOI.org, which powers Unpaywall, to its access systems so that when users search for an article an OA version will appear to them in the result lists if there's one available. This kind of service is of interest not only in terms of discoverability, but also as infrastructure that can support change in the scholarly publishing market.

PART II: OPEN ACCESS PRACTICES AT MIT

Publications

MIT faculty, researchers, and students produce thousands of publications each year in the form of journal and conference articles,[17] working papers, technical reports, theses,[18] books, and more. In the spirit of MIT's mission of disseminating knowledge, many of these publications are shared openly.

Journal Articles and Conference Papers

Open Access and Funder Policies

In March 2009, MIT faculty adopted one of the first open access policies in North America. Under the policy, faculty members grant to MIT non-exclusive permission to make their scholarly articles openly available. The policy calls on faculty to give the final, peer-reviewed versions of their articles to the MIT Libraries, whose staff put them in MIT's institutional repository (currently DSpace@MIT). Under MIT's policy, authors retain copyright in their scholarly works.

The MIT policy defines "scholarly article" per the 2002 Budapest definition, as literature describing the fruits of research that authors give to the world for the sake of inquiry and knowledge without expectation of payment. This literature is usually presented in peer-reviewed scholarly journals and conference proceedings.[19] The policy targets the "author's final manuscript"[20] of articles rather than the publisher's version.

Further inspired by the Budapest definition of open access, the faculty creators of the OA policy decided to share scholarly articles in DSpace under Creative Commons licenses. MIT's OA-policy articles use the Creative Commons Attributions-Noncommercial-Share Alike license,[21] which lets anyone share, excerpt, and create new materials from articles as long as they cite the original authors, their uses are not for profit, and they release derivative works under the same license.

Libraries staff manage the OA policy: They gather faculty-authored papers from department, lab, and center (DLC) websites and other repositories (such as PubMed Central [PMC] and arXiv), request papers from faculty authors, negotiate agreements with publishers to automatically receive papers (e.g., Springer, which sends papers directly into DSpace), and deposit and catalog the papers in DSpace@MIT.

As of fiscal year 2018, 43% of faculty-authored scholarly papers published since 2009 are openly available in DSpace through these methods.[22] The largest proportion (48%) of these OA papers come from publishers, who either directly send articles to the repository via the SWORD protocol, put papers out on their own websites under Creative Commons licenses,[23] or allow Libraries staff to deposit them to DSpace. Articles in

---

[17] From 2009 to 2017, MIT faculty published an average of 5,300 articles annually.

[18] The Institute Archives houses MIT theses, many of which (and all since 2004) are available online in DSpace@MIT.

[19] The MIT OA policy FAQ further notes: "Many written products are not encompassed under this specific notion of scholarly article, such as books, popular articles, commissioned articles, fiction and poetry, encyclopedia entries, ephemeral writings, lecture notes, lecture videos, or other copyrighted works. The Open Access Policy is not meant to address these kinds of works."

[20] This is the peer-reviewed, accepted version not yet with the publisher's formatting.

[21] About half of the articles in the Open Access Articles collection of DSpace@MIT are deposited under the OA policy and thus are shared under the CC BY-NC-SA license; the others are all rights reserved or are shared under other CC licenses.

[22] As of August 2018, there were more than 29,600 articles in the OA collection of DSpace.

[23] CC-licensed papers are either in all–open access journals or subscription journals with a paid OA option (so-called "hybrid" journals).

DSpace have permanent URLs, rich metadata that enable discovery, and a long-term commitment from the MIT Libraries to host and make them available.

About 37% of all OA papers in DSpace collected under the faculty OA policy have been found either on an MIT DLC site or an outside repository or website. Of these, 15% come from the Libraries' outreach—that is, sending emails to faculty once every 12–18 months to request papers that staff could not find another way.

Differences in deposit rates across the MIT DLCs are likely due in large part to differences in disciplinary culture and publishing practices. For example, 75% of physics faculty papers published between 2009 and 2017 are in DSpace, the highest proportion of any MIT department. This reflects physicists' long history of sharing different versions of papers among themselves and in the arXiv repository. Publishers of physics journals generally align with this culture of openness: The American Physical Society, for example, sends the final published version of papers directly to DSpace@MIT. (These direct deposits contribute to the high percentage of physics papers available.)

In other disciplines, publishers have shown significant resistance to open access publishing options, especially those options governed by institutional policies and implemented via institutional repositories. This affects faculty sharing. For example, only 26% of materials science and engineering papers are available in DSpace@MIT, primarily because some of the most popular journals in this discipline are published by Elsevier and the American Chemical Society, both of which require that researchers opt out of open access policies[24] and lock down papers in other ways.

There are also faculty at MIT who prefer to make open only the publisher's version of articles rather than the manuscript version the OA policy calls for. At the MIT Sloan School of Management, for example (for which 38% of faculty papers are in DSpace@MIT under the OA policy), some faculty choose to not share papers at all if they cannot share the publisher versions.[25]

In April 2017, MIT expanded the faculty open access policy to include MIT students, postdocs, and staff, who can now sign a voluntary, opt-in OA license. Nearly 900 people have signed.[26]

As with the vast majority of faculty open access policies across the country, MIT's policies allow authors to easily opt out for individual articles. By contrast, many funders have adopted policies that mandate open access, without waiver options. Federally funded MIT authors[27] are required to make articles and, in most cases, data, openly available under public access policies that came out of a 2013 directive from the White House Office of Science and Technology Policy. About a dozen federal agencies currently have policies under which authors must make their accepted manuscripts available within 12 months in a repository of the agency's choosing. With the exception of NIH, which has had public access requirements since 2008, compliance efforts are just getting underway.

---

[24] Since 2009, MIT faculty have opted out of (waived) the OA policy for nearly 1,500 papers, or 3% of articles they published. Researchers most commonly give one reason for opting out of the policy: that a publisher requires it (including Nature, PNAS, and Science). In some cases, the publisher's author agreement requires authors to opt out of the OA policy (e.g. Elsevier), and these papers are handled as if there has been a formal "opt out," though no waiver has been issued, so they are not counted as part of the 3%.

[25] Communications staff at Sloan, who build personal websites for faculty and gather papers to put on them, have repeatedly run into this issue as they try to balance publishers' policies against faculty desires.

[26] Affiliations of signers include: 409 graduate students; 110 researchers; 92 postdocs; 52 undergraduate students; 52 Libraries staff; 51 administrators/staff; 26 instructors/professors of the practice/emeriti.

[27] Nearly two-thirds of MIT's supported research is funded by federal agencies, including the National Institutes of Health, National Science Foundation, Department of Energy, and Department of Defense.

Sharing via Department, Lab, and Center Websites and Outside Repositories

Long before adopting the faculty OA policy, many MIT authors shared their research papers openly on the web. Libraries staff examined the MIT web domain as a source of journal and conference papers with faculty authors in 2010, as outreach under the OA policy got underway. They looked at almost two-thirds of DLCs, including 29 of 33 departments, and found that 35% of DLCs offered links to locally hosted papers, mostly via individual researcher pages[28] or lab sites. Nearly 60% of these papers were a version that fit under the OA policy—i.e., the author's final manuscript—and could be deposited into DSpace, where, as noted above, they are curated and made discoverable by Libraries staff. As of June 2018, 16% of papers deposited to DSpace under MIT's faculty OA policy came from DLC websites.[29]

Many researchers also post papers in subject-specific repositories, particularly arXiv and PMC. About 6% of all papers gathered under MIT's OA policy have come from arXiv; and 13% are from PMC, which is where NIH-funded authors are required to deposit manuscript versions of papers describing NIH-funded research.

Open Access Article Publishing

While there are many ways to support open access publication of journal articles, perhaps the most well known is the article processing charge (APC) model. At MIT, authors are increasingly taking advantage of the MIT Open Access Article Publication Subvention fund, which was launched by the Libraries in 2010 as a "last resort" reimbursement of up to $1,000 in processing fees per article in eligible all-open access, peer-reviewed journals. Between 2010 and March 2018, the Libraries paid $60,000 for about 100 articles from 70 applicants.[30] The most common publishers for which researchers request support are Public Library of Science (PLoS), Nature, Frontiers, and the American Physical Society.

MIT researchers are also increasing their OA journal publishing more generally. Between 2006 and 2016, the proportion of papers MIT authors published in all-open access journals grew from 2% to 15%[31] (higher than the increase in OA publishing globally). In 2016, the most popular OA journals for MIT authors were published by Nature (*Scientific Reports*, *Nature Communications*), Springer (*Journal of High Energy Physics*, via SCOAP3), and the American Physical Society (two physics journals via SCOAP3). The MIT Press currently publishes eight OA journals, primarily in the sciences, and frequently includes articles by MIT faculty. The *OA Journal of Design and Science* (*JoDS*) is a joint venture of the MIT Media Lab and the MIT Press.

Working Papers and Technical Reports

In some disciplines, particularly subsets of economics and management, researchers commonly produce working papers, which are often not peer reviewed and are not targeted under the MIT OA policies as they are currently implemented. Some DLCs use DSpace[32] for these papers (e.g., the Department of Economics[33]); others post on their websites. One example is Sloan's Center for Information Systems Research (CISR), which has a staff member who manages publications on their site. Visitors must register to view these

---

[28] Many faculty papers are found on their coauthors' MIT web pages (e.g., those of graduate students and postdocs).

[29] Examples of DLCs that post researchers' journal articles, conference papers, or book chapters on their websites include the Synthetic Neurobiology Group and the Senseable City Lab.

[30] There were 12 requests for the OA fund in 2015, 30 in 2016, 22 in 2017, and 20 in the first six months of 2018.

[31] From "Analysis of the international journal publishing activities for Massachusetts Institute of Technology with special emphasis on gold open access publishing," February 2018, by the Big Data Analytics Group in the Max Planck Digital Library.

[32] Nearly 12,000 of 99,000 items in DSpace are labeled as working papers or technical reports.

[33] There are 1,293 items in the Department of Economics Working Papers Series in DSpace, dating from 1967 to 2014. There appear to be no rights statements, which defaults to all rights reserved.

publications.[34] Other DLCs with working papers on their websites include MIT's Energy Initiative and the School of Architecture and Planning's Center for Real Estate.[35]

MIT authors are also heavy users of the Social Science Research Network (SSRN) repository, which houses working papers. Sloan, for example, has more than 2,600 papers from 289 authors in the repository, and MIT-authored papers have been downloaded more than 1.5 million times.

Books

Monographs are more difficult to track than journal articles; book publishers apply metadata inconsistently and publication databases (like Scopus and Web of Science) have historically had scant book coverage. That said, book vendor data indicates that in fiscal years 2017 and 2018, MIT faculty published 119 books, including textbooks, handbooks, and trade books.[36] Of those, 38 were published by university presses (including 16 by The MIT Press[37]); four by societies; and 77 by commercial presses.

The primary focus of early open access efforts was journal articles, but there are several promising projects focusing on open access scholarly monographs. The MIT Press has been a leader in open access book publishing for over 20 years, starting with the 1995 publication of *City of Bits* and continuing with a steady stream of new open access books every year. Beyond MIT, other notable open access book projects include the Association of Research Libraries' TOME (Toward an Open Monograph Ecosystem) and Luminos from the University of California Press. While we do not know what proportion of MIT-authored books are open access, there have been significant commitments by individual faculty to make their books openly available.[38]

In addition, there is considerable support on many campuses and at the federal and state level for open access textbooks, including a recent $5 million open textbook grant program funded by Congress.

Data

MIT researchers have data-sharing requirements from both funders and journals. The 2013 White House directive (described in Part I) requires that data resulting from unclassified research supported by federal funding "be stored and publicly accessible to search, retrieve, and analyze."[39] Funders are generally less explicit with requirements related to data than they are about articles in terms of which repositories authors should use to make data publicly accessible.

Some journals have their own data-sharing policies, which range from *Science*'s policy that "reasonable requests must be fulfilled," to public repository mandates such as those required by some Wiley journals and the SpringerNature journal *Genome Biology*.

Though MIT Libraries and MIT's Information Systems & Technology (IS&T) staff support researchers in creating data management plans to meet funder requirements, finding repositories for data storage, and using tools like LabArchives and Dropbox, data management and sharing practices across the Institute are

---

[34] The CISR license statement says, "Downloaded publications may be reproduced free of charge for educational purposes, provided the copyright statement appears on each copy. Please note that these documents may not be copied or distributed for commercial purposes."

[35] A sampling of MITEI's papers shows that copyright belongs to MIT with all rights reserved. A sampling of Center for Real Estate papers included no rights statements, which means they default to all rights reserved.

[36] Data comes from the print and electronic book company GOBI.

[37] Four of The MIT Press books are open access.

[38] Two examples are Professor Eric von Hippel's open access versions of his MIT Press books, and Professor John Lienhard's open sharing of *A Heat Transfer Textbook*.

[39] The directive defines research data as "the recorded factual material commonly accepted in the [research] community as necessary to validate research findings."

varied and decentralized. There is no comprehensive record of where MIT research data is stored, published, and/or shared.

That said, the Libraries has gathered some information about these practices in surveys. Results from a 2017 MIT Libraries survey show that 58% of respondents share their "research outputs" (including articles and data) through Dropbox and 51% through Google Drive. Meanwhile, 11% or fewer share these outputs through DSpace[40] or another repository, such as PubMed Central or Dataverse.

Many MIT DLCs store and share research data on local servers. These include MIT's Computer Science & Artificial Intelligence Lab (CSAIL), which offers dataset publishing and storage; the Trancik Lab, of Institute for Data, Systems, and Society faculty member Jessika Trancik; and the Living Wage Calculator from Department of Urban Studies and Planning faculty member Amy Glasmeier.

Well-run subject repositories exist for some research data, and researchers in those fields are well served. Examples include the Inter-university Consortium for Political and Social Research, Protein Data Bank, and Data Dryad. However, the majority of disciplines do not have fully developed repositories, and the Libraries data management team frequently hears from researchers in engineering and humanities fields,[41] for example, who would like to store their data through MIT. These researchers do not always find MIT-supported options sufficient.

Code

As with data, it is difficult to generalize about practices with MIT-created software and code—and perhaps even more so because there are no federal funder public access mandates for code, nor do most publishers and journals[42] require it be made openly available in designated repositories.

One of the most popular code repositories is GitHub. IS&T supports an MIT GitHub that hosts more than 25,500 project repositories, 18% of which are public.[43] But there's no requirement to use MIT GitHub, so it is not possible to know how many MIT staff and researchers have created their own projects elsewhere on the site.

There is some use data about code in the MIT Libraries' 2017 survey: 28% of MIT students, 18% of faculty, and 25% of researcher respondents use GitHub to share research outputs (including code).[44] As with publications and data, researchers also host code on their DLC websites, as is the case with Sloan professor Andrew Lo's website and the Open Agriculture Initiative in the Media Lab.

In terms of licensing code, MIT's Technology Licensing Office (TLO) offers technology disclosure forms and support for software: one form for software that "incorporates a patentable process or algorithm," and one for software that "encompasses the copyrightable software code." Though there is no formal requirement to disclose software, it is often important to do so if software is developed under sponsored research. And the TLO particularly recommends disclosure if people "believe their research could be commercialized for public use and benefit."

---

[40] In FY2017, for example, Libraries staff created the Data Related to Publications collection in DSpace, and these records are often linked to articles collected under the OA policy.

[41] The research data in these cases may often be long-tail data. This kind of data varies "radically in source, discipline, size, subject, provenance, funding, format, longevity, location and complexity, [and] are less likely to adhere to common standards" according to the Research Data Alliance.

[42] Some exceptions: The data journal *GigaScience* requires code used in its papers to be available. The publisher Nature encourages code sharing, and the SpringerNature title *Genome Biology* requires it.

[43] From MIT Office of the Vice President for Open Learning.

[44] According to GitHub, the most popular open-source license used on the platform is the MIT License. Others are popular too. This MIT-created GitHub repository uses BSD.

About 10% of 800 annual disclosure requests to the TLO are for non-patented software. Between 40% and 60% of these software disclosures are open-source requests. The licenses vary, though TLO encourages most people to use either the BSD or GPLv2 license.

According to Daniel Dardani, MIT's chief software and algorithms technology licensing officer, "We aim to be pragmatic in our approach, but as a general rule, the TLO avoids using open-source licenses with expressed patent grant language because of the potential conflict they pose given MIT's substantial and diverse portfolio of patented technologies, many of which are exclusively licensed to companies."

Educational Materials

Stellar

As of April 2018, there were 944 class sites with materials on the course management system Stellar.[45] (This number includes classes on Learning Modules, a newer site also run by IS&T that is integrated with Stellar; instructors can choose to use one or the other or a hybrid of the two.[46]) Stellar sites are by default open only to class members, though instructors can change this setting to allow anyone at MIT, or individual non-MIT users to view materials. Of the 944 Stellar sites, 74% are open to the class only, 19% are open to the MIT community, and 7% are open to the public.

OpenCourseWare

As of February 2018, there were 2,489 courses live on the OpenCourseWare site.[47] All OCW content is available exclusively under a CC BY-NC-SA license. (There is some third-party material on the site; it is marked as being used under fair use and is not available under CC licenses.)

MITx

As of February 2018, there were approximately 140 "unique"[48] courses on MITx, MIT's edX site, where anyone can register for and view content. However, with the exception of about five courses with materials under Creative Commons licenses, all courses and materials available on MITx are marked all rights reserved.[49] Procedurally, faculty and course teams do not transfer copyright to MIT; they sign a license so MITx can use it.[50]

Other Websites and Platforms

As they do with publications, data, and code, some DLCs post open learning materials on their websites.[51] Examples include:

- Sloan LearningEdge, modeled after OCW, which includes case studies and other tools developed by Sloan faculty and students and under Creative Commons licenses.

---

[45] Stellar is not used universally; Sloan has started using the course management program Canvas.

[46] Learning Modules was initially scheduled to fully replace Stellar in 2016–2017, but this is on hold for the "foreseeable future."

[47] OCW materials appear in other venues, including YouTube and the OER Commons.

[48] "Unique" means it has one associated course number, but each course can then break down into different modules. In the 140 courses there were 540 modules with materials including text, videos, and articles.

[49] MITx uses edX user content agreements and terms of service.

[50] Some faculty have posted courses without signing a license agreement because they do not want MIT to sublicense their material.

[51] There are also non-open (and for-pay) learning materials; e.g., MITx Pro and MicroMasters.

- The MIT Case Study Initiative, launched in academic year 2017 by School of Architecture and Planning faculty and a team of developers, to feature customizable case studies with video, images, data, documents, and assessments for urban development educators.

- MIT Blossoms, which offers science and math videos for high school teachers, generally under Creative Commons licenses.

There are also learning materials hosted on YouTube, where MIT-affiliated channels offer nearly 440 videos, including videos from OCW, the MIT Alumni Association, the School of Engineering, the Department of Mechanical Engineering, CSAIL, Sloan, and others. (Other MIT-affiliated content is elsewhere on YouTube). MIT videos on YouTube are either under the "Standard YouTube License" (all rights reserved by content owners) or Creative Commons licenses.

CONCLUSION

In 2013, then-Director of the MIT Libraries Ann Wolpert wrote: "There is no doubt that the public interests vested in funding agencies, universities, libraries, and authors, together with the power and reach of the Internet, have created a compelling and necessary momentum for open access. It won't be easy, and it won't be inexpensive, but it is only a matter of time."

As this white paper describes, significant progress has been made—globally, nationally, and locally— in providing open access to scholarly research since Wolpert's prediction. Building on the policies, infrastructure, and practices already in place, the MIT Ad Hoc Task Force on Open Access to MIT's Research will develop a draft set of recommendations to further the Institute's mission of disseminating research that can be brought to bear on the world's greatest challenges.

APPENDIX A. OPEN ACCESS TASK FORCE MEMBERSHIP

Co-chairs

Hal Abelson, *Class of 1922 Professor, Department of Electrical Engineering and Computer Science*

Chris Bourg, *Director of MIT Libraries*

Members

Peter Bebergal, *Technology Licensing Officer, Technology Licensing Office*

Robert Bond, *Chief Technology Officer, Lincoln Laboratory*

Herng Yi Cheng '18, *Department of Mathematics*

Isaac Chuang, *Professor of Electrical Engineering and Computer Science; Senior Associate Dean of Digital Learning*

Christopher Cummins, *Henry Dreyfus Professor of Chemistry, Department of Chemistry*

Deborah Fitzgerald, *Leverett Howell and William King Cutten Professor of the History of Technology, Program in Science, Technology, and Society*

Mark Jarzombek, *Professor, Department of Architecture*

Nick Lindsay, *Director of Journals and Open Access, The MIT Press*

Tom Pollard, *Postdoctoral Associate, Institute for Medical Engineering and Science*

Jack Reid G '18, *Technology and Policy and Aeronautics and Astronautics*

Karen Shirer, *Director of Research Development, Office of the Provost*

Bernhardt Trout, *Professor, Department of Chemical Engineering*

Matthew Vander Heiden, *Associate Professor, Department of Biology*

Eric von Hippel, *T Wilson (1953) Professor in Management, Sloan School of Management*

Jaren D. Wilcoxson, *Counsel, Office of the General Counsel*

Staff to the Committee

Ellen Finnie, *Head, Scholarly Communications and Collections Strategy, MIT Libraries*