OPEN ACCESS

PETER SUBER

OPEN ACCESS

PETER SUBER

The MIT Press | Cambridge, Massachusetts | London, England

SERIES FOREWORD

The MIT Press Essential Knowledge series presents short, accessible books on need-to-know subjects in a variety of fields. Written by leading thinkers, Essential Knowledge volumes deliver concise, expert overviews of topics ranging from the cultural and historical to the scientific and technical. In our information age, opinion, rationalization, and superficial descriptions are readily available. Much harder to come by are the principled understanding and foundational knowledge needed to inform our opinions and decisions. This series of beautifully produced, pocket-sized, soft-cover books provides in-depth, authoritative material on topics of current interest in a form accessible to non-experts. Instead of condensed versions of specialist texts, these books synthesize anew important subjects for a knowledgeable audience. For those who seek to enter a subject via its fundamentals, Essential Knowledge volumes deliver the understanding and insight needed to navigate a complex world.

Bruce Tidor
Professor of Biological Engineering and Computer Science
Massachusetts Institute of Technology

I have worked full-time for a decade to foster open access (OA) to science and scholarship. During that time I have often boiled down the big message into short talks and written long articles exploring small subtopics in detail. This book is an attempt at something in between: a succinct introduction to the basics, long enough to cover the major topics in reasonable detail and short enough for busy people to read.

I want busy people to read this book. OA benefits literally everyone, for the same reasons that research itself benefits literally everyone. OA performs this service by facilitating research and making the results more widely available and useful. It benefits researchers as readers by helping them find and retrieve the information they need, and it benefits researchers as authors by helping them reach readers who can apply, cite, and build on their work. OA benefits nonresearchers by accelerating research and all the goods that depend on research, such as new medicines, useful technologies, solved problems, informed decisions, improved policies, and beautiful understanding.

But OA only does this good work insofar as we actually implement it, and the people in a position to implement it tend to be busy. I'm thinking about researchers

themselves and policymakers at stakeholder institutions such as universities, libraries, publishers, scholarly societies, funding agencies, and governments.

My honest belief from experience in the trenches is that the largest obstacle to OA is misunderstanding. The largest cause of misunderstanding is lack of familiarity, and the largest cause of unfamiliarity is preoccupation. Everyone is busy. There has been organized opposition from some publishers, but that has been a minor impediment by comparison.

The best remedy to misunderstanding is a clear statement of the basics for busy people. Only some fellow specialists will wonder, with me, whether I've been too brief with some essential subtopics. But I knew that a larger book would miss the audience of busy people. Elaboration, documentation, research findings, case studies, and finergrained recommendations are available in the voluminous literature online (most of it OA), including my own articles (all of them OA).[1]

This book will itself be OA twelve months after it appears in print. (I'm glad you asked.) If you can't wait, everything I've said here I've said in some form or another in an OA article.

I have freely incorporated some relevant earlier writings into this book, improving on them when I could. Notes at the end of the book indicate which pieces I adapted or

incorporated into which sections. I chose this method as a solution to a pair of dilemmas. I did not want to hide the fact that I was making use of my previous work, but neither did I want to make any section into a stream of self-quotation and self-citation. I did not want to fail to benefit from my own previous work, but neither did I want to miss opportunities to clarify, update, or improve it.

This little book doesn't say much about kindred topics such as open data, open educational resources, open government, free and open-source software, or open science (combining OA texts, open data, and open-source software, and providing these sorts of openness at every stage of a research project, not just at the end in reporting results). Some of the kindred forms of scholarly openness might soon be covered by other volumes in this series.

Douglas Bennett, Len Clark, Darius Cuplinskas, Robert Darnton, Urs Gasser, Melissa Hagemann, Rick Johnson, Heather Joseph, Robert Kiley, Sue Kriegsman, Harlan Onsrud, John Palfrey, Lisbet Rausing, Stuart Shieber, David Skurnik, and Gigi Sohn.

I dedicate this book to the thousands of people in every field and country who have dedicated themselves to the realization of OA. The ones I know personally are already too numerous to thank by name in the preface to a short book, and the fact that there are more than I could thank by name—even if I tried—fills me with admiration, gratitude, and optimism.

CONTENTS

WHAT IS OPEN ACCESS?

Shifting from ink on paper to digital text suddenly allows us to make perfect copies of our work. Shifting from isolated computers to a globe-spanning network of connected computers suddenly allows us to share perfect copies of our work with a worldwide audience at essentially no cost. About thirty years ago this kind of free global sharing became something new under the sun. Before that, it would have sounded like a quixotic dream.

Digital technologies have created more than one revolution. Let's call this one the access revolution.

Why don't more authors take advantage of the access revolution to reach more readers? The answer is pretty clear. Authors who share their works in this way aren't selling them, and even authors with purposes higher than money depend on sales to make a living. Or at least they appreciate sales.

Let's sharpen the question, then, by putting to one side authors who want to sell their work. We can even acknowledge that we're putting aside the vast majority of authors.

Imagine a tribe of authors who write serious and useful work, and who follow a centuries-old custom of giving it away without charge. I don't mean a group of rich authors who don't need money. I mean a group of authors defined by their topics, genres, purposes, incentives, and institutional circumstances, not by their wealth. In fact, very few are wealthy. For now, it doesn't matter who these authors are, how rare they are, what they write, or why they follow this peculiar custom. It's enough to know that their employers pay them salaries, freeing them to give away their work, that they write for impact rather than money, and that they score career points when they make the kind of impact they hoped to make. Suppose that selling their work would actually harm their interests by shrinking their audience, reducing their impact, and distorting their professional goals by steering them toward popular topics and away from the specialized questions on which they are experts.

If authors like that exist, at least they should take advantage of the access revolution. The dream of global free access can be a reality for them, even if most other authors hope to earn royalties and feel obliged to sit out this particular revolution.

It's enough to know that their employers pay them salaries, freeing them to give away their work, that they write for impact rather than money, and that they score career points when they make the kind of impact they hoped to make.

These lucky authors are scholars, and the works they customarily write and publish without payment are peer-reviewed articles in scholarly journals. *Open access* is the name of the revolutionary kind of access these authors, unencumbered by a motive of financial gain, are free to provide to their readers.

> Open access (OA) literature is digital, online, free of charge, and free of most copyright and licensing restrictions.

We could call it "barrier-free" access, but that would emphasize the negative rather than the positive. In any case, we can be more specific about which access barriers OA removes.

A price tag is a significant access barrier. Most works with price tags are individually affordable. But when a scholar needs to read or consult hundreds of works for one research project, or when a library must provide access for thousands of faculty and students working on tens of thousands of topics, and when the volume of new work grows explosively every year, price barriers become insurmountable. The resulting access gaps harm authors by limiting their audience and impact, harm readers by limiting what they can retrieve and read, and thereby harm research from both directions. OA removes price barriers.

Copyright can also be a significant access barrier. If you have access to a work for reading but want to translate it into another language, distribute copies to colleagues, copy the text for mining with sophisticated software, or reformat it for reading with new technology, then you generally need the permission of the copyright holder. That makes sense when the author wants to sell the work and when the use you have in mind could undermine sales. But for research articles we're generally talking about authors from the special tribe who want to share their work as widely as possible. Even these authors, however, tend to transfer their copyrights to intermediaries—publishers—who want to sell their work. As a result, users may be hampered in their research by barriers erected to serve intermediaries rather than authors. In addition, replacing user freedom with permission-seeking harms research authors by limiting the usefulness of their work, harms research readers by limiting the uses they may make of works even when they have access, and thereby harms research from both directions. OA removes these permission barriers.

Removing price barriers means that readers are not limited by their own ability to pay, or by the budgets of the institutions where they may have library privileges. Removing permission barriers means that scholars are free to use or reuse literature for scholarly purposes. These purposes include reading and searching, but also redistributing, translating, text mining, migrating to new media,

Terminology

When we need to, we can be more specific about access vehicles and access barriers. In the jargon, OA delivered by journals is called *gold OA*, and OA delivered by repositories is called *green OA*. Work that is not open access, or that is available only for a price, is called *toll access* (TA). Over the years I've asked publishers for a neutral, nonpejorative and nonhonorific term for toll-access publishers, and *conventional publishers* is the suggestion I hear most often. While every kind of OA removes price barriers, there are many different permission barriers we could remove if we wanted to. If we remove price barriers alone, we provide *gratis OA*, and if we remove at least some permission barriers as well, we provide *libre OA*. (Also see section 3.1 on green/gold and section 3.3 on gratis/libre.)

long-term archiving, and innumerable new forms of research, analysis, and processing we haven't yet imagined. OA makes work more useful in both ways, by making it available to more people who can put it to use, and by freeing those people to use and reuse it.

OA was defined in three influential public statements: the Budapest Open Access Initiative (February 2002), the Bethesda Statement on Open Access Publishing (June 2003), and the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities (October 2003).[1] I sometimes refer to their overlap or common ground as the BBB definition of OA. My definition here is the BBB definition reduced to its essential elements and refined with some post-BBB terminology (green, gold, gratis, libre) for speaking precisely about subspecies of OA. Here's how the Budapest statement defined OA:

> There are many degrees and kinds of wider and easier access to [research] literature. By "open access" to this literature, we mean its free availability on the public internet, permitting any users 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. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited.

Here's how the Bethesda and Berlin statements put it: For a work to be OA, the copyright holder must consent in advance to let users "copy, use, distribute, transmit and display the work publicly and to make and distribute derivative works, in any digital medium for any responsible purpose, subject to proper attribution of authorship."

Note that all three legs of the BBB definition go beyond removing price barriers to removing permission barriers, or beyond gratis OA to libre OA. But at the same time, all three allow at least one limit on user freedom: an obligation to attribute the work to the author. The purpose of OA is to remove barriers to all legitimate scholarly uses for scholarly literature, but there's no legitimate scholarly purpose in suppressing attribution to the texts we use. (That's why my shorthand definition says that OA literature is free of "most" rather than "all" copyright and licensing restrictions.)

The basic idea of OA is simple: Make research literature available online without price barriers and without most permission barriers. Even the implementation is simple enough that the volume of peer-reviewed OA literature and the number of institutions providing it have grown at an increasing rate for more than a decade. If there are complexities, they lie in the transition from where we are now to a world in which OA is the default for new research. This is complicated because the major obstacles