



Guide ■ Top Challenges for Journal Publishers:

A Management Guide to Workflow and Keeping Up With the Industry

Table of Contents

Introduction	2
Chapter 1: Authoring and Editing	4
Chapter 2: Production	7
Chapter 3: Distribution	9
Chapter 4: Digital Asset Management and Archiving	13
Chapter 5: Partners	16
The Takeaway	20



Introduction

Last year, the journal publishing industry marked a milestone: 350 years since the publication of the first scientific journal, *Philosophical Transactions of the Royal Society of London*, in 1665.

The goal of that publication’s editor—to publish scientific reports, book reviews, obituaries, and law reports—was not that far from the aims of many journals today.

And what a journey it has been. Estimates vary, but by one calculation, there are **more than 28,100** active scholarly peer-reviewed journals, and **in his industry-standard 2015 STM Report**, Mark Ware puts English-language journal industry revenue at about \$10 billion in 2013, as part of a broader STM information publishing market worth more than \$25 billion. Most of the revenue comes from the United States. And it’s increasingly profitable.

The Challenges

The mission of journal publishers hasn’t changed. So what *is* changing?

Like all businesses in the era of the internet, the ground is shaking under the journal publishing industry. Customers mostly want the same product, the journal article. But they want to access these articles in new places, in different formats, at lower prices. And they want it faster.



Customers want to access these articles in new places, in different formats, at lower prices. And they want it faster.



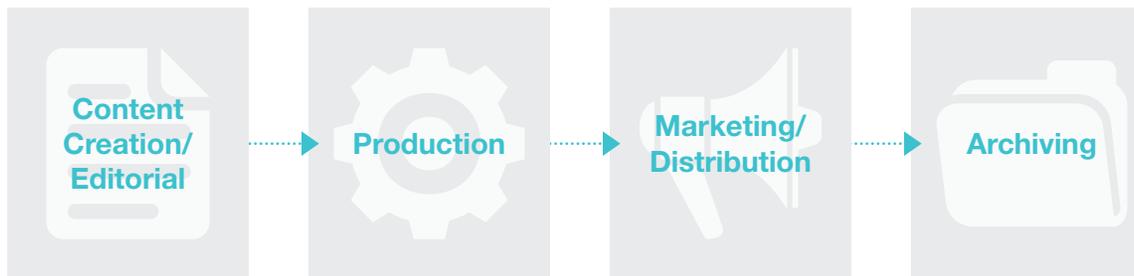
The dramatic rise in subscription prices (three to four times faster than inflation since 1986) reshaped the journal publishing business and put pressure on all of the industry participants to become more efficient and more productive.

Today's mantra is Quality + Velocity. Do what you're doing, but do it faster. And, of course, never sacrifice quality.

As publishers apply this dictum to their publishing workflows, they will discover unexpected opportunities for improvement. They may be able to implement some of those improvements internally. Others will require outside expertise.

The Forward Flow

Journal Publishing Workflow



These categories simply and broadly define the major steps in the journal publishing process. They're easy to identify from the 10,000-foot view, but on the ground, the steps are often redundant, inefficient, and error-prone.

In this guide, we'll pause on each category in our publishing workflow journey.

We'll help publishers identify key areas where problems arise. We'll offer tips on how to best analyze whether the problems are within your organization, and we'll offer insights into how you might address issues in-house or with outside resources.

Let's look at these opportunities to improve workflows in more detail. We can start with authors, scientists, and researchers. Without them, there's nothing to publish.



Today's mantra is Quality + Velocity. Do what you're doing, but do it faster. And, of course, never sacrifice quality.



Chapter 1: Authoring and Editing

Top Challenges: Journals' workflows—from submission through proofing cycles—are interrupted because of authors' and publishing house editors' lack of expertise in authoring and editing software. Also, editors struggle to manage cadres of both authors and peer-reviewers.

Authors and editors are true partners in the journal publishing process, which has become so complex. New CRediT metadata differentiates authors' roles, such as who was responsible for analysis, experimental design, data collection, and of course, the actual writing. Neither authors nor editors work in a solitary way.

Editors face day-to-day challenges supporting scholarly content creators and their workflows. Part of the problem stems from the authors being scientists first and authors second. Article content demands pinpoint accuracy, built on a foundation of XML structure.

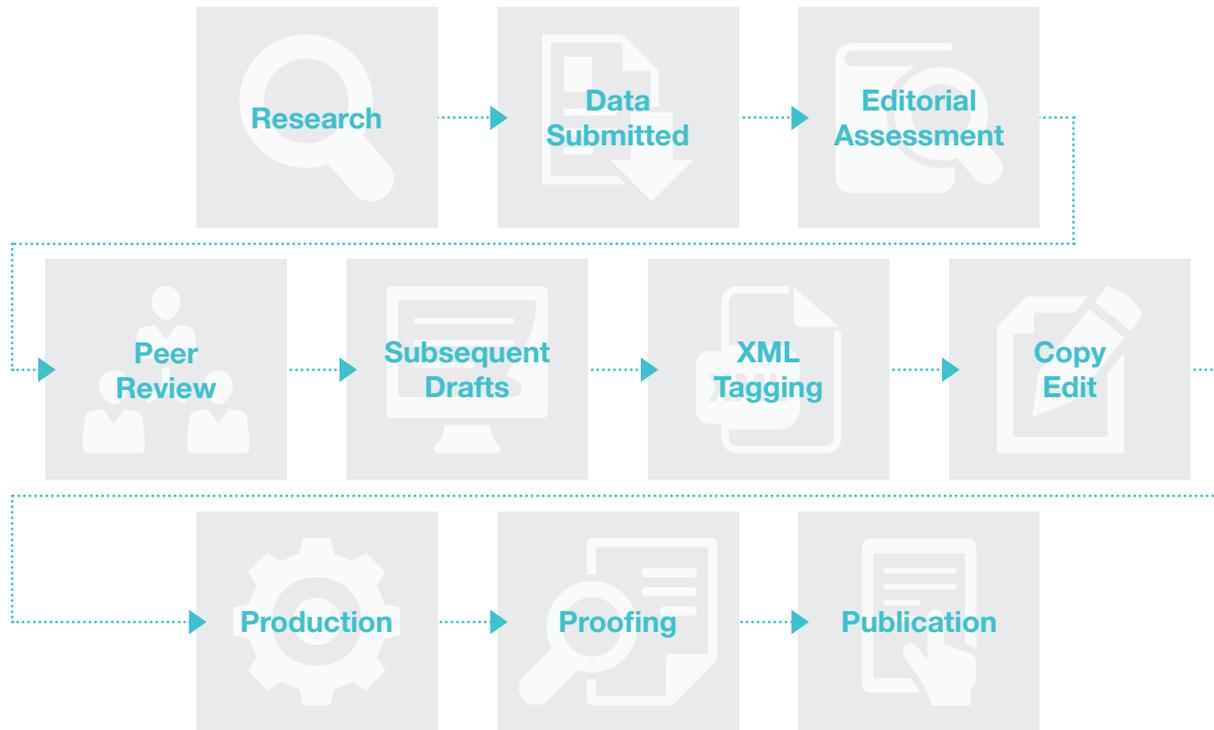
Journal authors link to editors as soon as an article is considered for publication. Online submission and peer-review systems control the flow. Text conversion, editing, and proofing systems hone the output.



Editors face day-to-day challenges supporting scholarly content creators and their workflows.



Authoring & Editorial Workflow



The Submissions Process

Article submission standards for scholarly publishing can be as rigorous as the science itself. Most journals dictate strict requirements for formatting, including references, keywords, and supplemental material. *JAMA*, the *Journal of the American Medical Association*, for example, **devotes 25,000 words to “Instructions for Authors.”** Elsevier’s *American Journal of Preventive Medicine* has a **16-page “Author Instructions” document.**

At the other end of the spectrum, publisher Elsevier now allows authors to submit Microsoft Word or PDF files without rigid formatting structures. It saves time for authors and gives them greater flexibility when making multiple submissions. The increasing automation of editorial functions support a more relaxed submission process.



Editorial Metrics

Improving the editorial workflow calls for an emphasis on measurement. Here’s what can be tracked:

- The time from submission to the first decision on suitability
- Overall acceptance and rejection rate
- The time needed for authors to revise
- Changes and who is responsible for them
- The time from full acceptance to publication

Much of the work takes place within each publishing group, but outside consultants can provide valuable insights into the state-of-the-art editorial management and production systems.



Editorial Reinigorated

Editorial is by nature collaborative. It's a partnership among authors, peer-reviewers, the editor, and the journal production team. Because of the core importance of peer review, the editor's function is sometimes more like traffic cop than a connoisseur of language.

Adding to the challenge, the number of authors per article is 6.57, according to the National Library of Medicine. At the STM Digital Publishing seminar late last year, a speaker shared an article that had 1,268 authors, taking up the first four pages in print. **According to an article in *The Winnower***, by 2034, there will be an average of 8 authors per paper.



Editorial is by nature collaborative. It's a partnership among authors, peer-reviewers, the editor, and the journal production team.

Article tracking software provides an interface among authors, peer-reviewers, and editors. These systems have streamlined a cumbersome manual process.

It's interesting to note that even with these systems, the time from submission to acceptance of the average scholarly paper **is still roughly 100 days**—as it has been for more than 30 years. Is the science more complex? Are the reviewers more exacting? No one seems to know, but shrinking this timeframe is the essence of improved process for publishers.

In the End

When the new issue has “gone to bed,” the editor's role is not over. Perhaps the largest task still looms: making sure there will be a next issue. **As Ronan Cormacain puts it**, “My nagging worry is that there will not be enough articles for the next issue—I am sure that all but the most successful of journals share this worry.”

Meanwhile, the text moves on to production, where another group of technologist specialists takes on the publishing responsibility.



Content Beyond Words

Content creation now covers a much broader spectrum, from the humble 140-character tweet (sometimes called “scholarly microblogging”) to elaborate videos or datasets that can accompany a journal article. The Force 11 Data Citation Principles have been endorsed by more than 100 organizations, including large publishers like Elsevier and PLOS.

The American Meteorological Society now encourages authors to include animations, small datasets, computer codes, and oversized tables. The Association for Computing Machinery promotes animations in research articles, while Canadian Science Publishing says “Don't be shy—create a video abstract.” Elsevier's Content Innovation portfolio includes 3-D images, interactive graphics, podcasts, and videos.





Chapter 2: Production

Top Challenges: Journals struggle with making the best choices to maximize impact as production needs to support many outputs, print and digital.

XML, JATS and MathML. HTML5, XHTML and CSS. JPG, GIF, and SVG. EPUB 3. This alphabet soup evokes the complexity of journal production today. These standards and structures are powerful and complex. They *are* powerful because they are complex; complexity is the tradeoff for power.

There's an old saying, "Faster, better cheaper: Choose two, you can't have all three." But for journals today, it has to be the trio.

Faster: The push is on to make science available sooner; however, the peer-review cycle is stuck in time, resistant to efficiency. That leaves editorial and production. Journals compete, and the speed of publication sways author choices. Some STM publishers receive feedback from authors that the time it takes for articles to be published online—up to five or six weeks—is unacceptably slow.

Better: Journals must create content that is consistently clear, reliable, authoritative, and relevant.

Cheaper: Management is unrelenting: There's always more blood to be found in the stone.

There's an old saying, "Faster, better cheaper: Choose two, you can't have all three." But for journals today, it has to be the trio.



Pity the PDF

Most journal readers **still prefer PDFs**. The traditional article structure is well-suited to PDF's rigidity. Are there better ways to present a journal article? Perhaps. But for now, the PDF is primary.

PDF's rigid boundaries are also its great liability. Letter and A4 pages don't work at smaller sizes. Scrolling isn't efficient. Accessibility is compromised.

Fortunately, it's now possible to generate PDF files dynamically using XML-tagged content. These more fluid structures augur well for multichannel publishing.

At the same time, developers cook up ways to surround PDFs with greater functionality. They're now routinely displayed within web browsers, making it more convenient for functions such as advanced search, annotation, linking, and sharing.

Journal publishing will still be working with PDFs for some time to come.

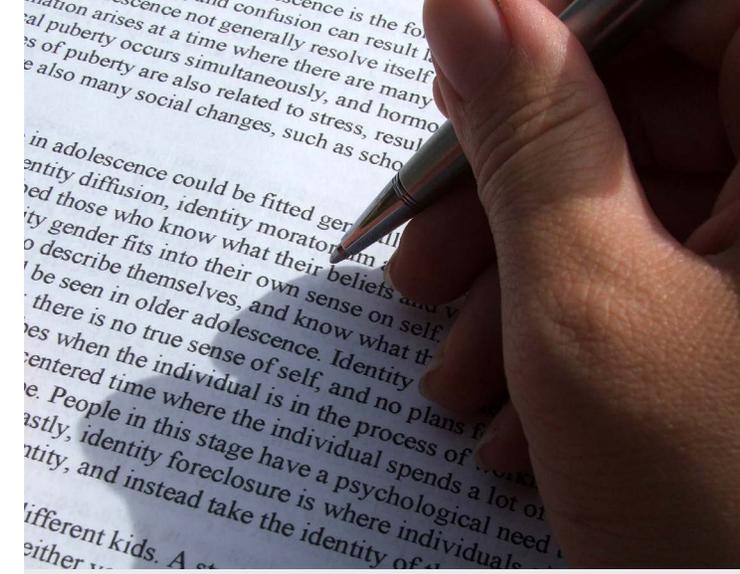
Enter Standards

Every dollar invested in increased support of industry standards within a publishing workflow is a dollar that multiplies across the three checkpoints of faster, better, cheaper. Custom code, by comparison, is \$10 wasted.

Indeed, like all technology, the sands of standards constantly shift. For example, is it time to jump from NLM 2.3 to JATS? You need outside experts and collegial advice to develop answers to changing standards. Backwards compatibility is key: If the standard you're moving toward can still support its aging granddad, it's probably a standard worth supporting.

And So to Press

Tagged. Edited. Composed. Re-edited. Author corrections. New page proofs. And so to press. Whatever press might mean in these days—distribution is just a smartphone away.



Multichannel Publishing

Changing priorities for multiple media output is at the heart of journal production today.

Formats are changing too quickly—and too dramatically—to manage successfully. There are many side roads that lead to dead ends. Yet the pressure is relentless to deliver for print, for aggregators, for mobile, for accessibility, with all things digital.



Decoding XML

XML has been on a long journey through the world of scholarly publishing, that began with its predecessor, SGML, launched in the 1970s.

Publishers welcomed XML as SGML's progeny, in part because it was simpler than SGML. And so complexity is XML's birthright and an overhead it still carries today. XML links to a family of related standards, including XSLT, XHTML, CSS, MathML, RELAX NG and the Journal Article Tag Suite (JATS).

Common XML-based models like JATS are an invaluable framework, but they are not specifications. They're designed to accommodate the needs of thousands of publishers; a specific publisher needs to specify how its JATS XML should be tagged.

It's nearly 20 years since the approval of the XML 1.0 spec and perhaps the most controversial question remains: Can we get authors to use it? The short answer: Yes, if we hide it from them, so they don't realize they're using it.

XML's complexity signals its power. It's a markup language that can describe both format—how content is intended to look— and structure and semantics, a view of the underlying “meaning” of the content.

Where the early benefit of XML structures was savings on composition, the legacy benefit is semantic discoverability. The wide adoption of JATS for journals means a standard interoperable semantic structure, enabling a new world of journal discoverability.

Chapter 3: Distribution

Top Challenge: Journals, like other publishers, are trying to keep up with the move to mobile.

“Print-only.” How quaint that term seems in the new era of scholarly journal distribution. One prominent scholarly journal distributor offers its clients three formats: electronic only; electronic + print; and print only. Even three formats is beginning to sound retro.

As Michael Clarke points out, journals transitioned to online delivery rapidly, more rapidly than books. The real driver was search and discovery. Every scientist wants to lay alpha claim to a new line of thinking and most would rather not spend time repeating an experiment someone else conducted. Search and discovery help these efforts.

Another reason is that the scholarly publishing industry was already heavily invested in XML (and its predecessor SGML). As the demand for online access grew, it was easy for publishers to convert their tagged content to a variety of online formats.



Journals transitioned to online delivery rapidly, more rapidly than books. The real driver was search and discovery.



Discovery

With near-universal digitization, distribution today is more about discovery than it is about formats. Being read is far more important for researchers than any other single factor motivating researchers to publish. Not to be read is the equivalent of never having been published.

Readers access journals via so many portals, from institutional databases to aggregators, from publishers' and authors' sites. Each of these needs to be factored into the journal publisher's discoverability strategy. Search is central to discovery, both Google Search and Google Scholar.

The preeminent study on the topic, *How Readers Discover Content in Scholarly Publications*, illustrates the multitude of paths for reader "navigation and discovery."

Google Scholar is having a major effect on article discoverability. **A recent research project noted that**, because of search, finding older articles is nearly as easy as finding recently published articles. The study showed citations to articles that are at least 10 years old had grown 28 percent between 1990 and 2013. Further, Google Scholar can **retrieve twice as many relevant articles** as the respected PubMed database.



With near-universal digitization, distribution today is more about discovery than it is about formats.



Metadata 101

Metadata is the most complex *simple* idea in all of publishing. It is too broadly defined, constantly in transition, and loosely used in practice. **OCLC's Chip Nilges provides a practical definition:** "Metadata is the language used to communicate information about a book or piece of content necessary for discovery, access, and sales."

Where it's found: Google and other general search engines, Google Scholar and other dedicated search platforms, content aggregator sites, in the library, and other places. Without metadata, published research is all but invisible. Metadata makes things known.

When it comes to standards, metadata has too many. There's the venerable Dublin Core, a simple set of elements designed to describe documents (and other objects) on the internet. For the book supply chain, there's ONIX. In the library world, there are MARC records. There's the still-new (2011) schema.org for semantic content in browsers. Knowing which standards matter and how to use them calls for expert advice. For example, Google Scholar has **very specific rules for inclusion.**

Bill Kasdorf captured metadata's essential problem when he **coined the term "messadata"**—too much metadata inconsistently conveyed, incorrect data, shifting standards. Unquestionably metadata remains a specialty—the average journal editor cannot be expected to remain fully up to speed. Larger organizations can afford full-time metadata specialists. For the rest, outside knowledge resources are essential and smart and experienced consultants are available for hire.

Discovery, of course, regardless of the platform, is inextricably linked to metadata.

The goal of distribution then is to be where a reader is searching at that moment. If readers have to jump through hoops to access an article, they will increasingly choose to move on to the next article on the same subject.

Perceived Value of Scholarly Articles, by Carol Tenopir *et al.*, demonstrated that other than the article topic, having the article available online at no personal cost to the reader was the second most prevalent factor influencing reader selection (the *source* of the article tied for second place).

Mobile and Scholarly Content

Access to scholarly articles via mobile **is a subject of ridicule.** This is not a good sign.

Audiences, even scholarly researchers, are moving to mobile faster than they moved from library-only access to internet access in the mid-1990s. The research site Asymco predicts 100 percent mobile adoption by 2020.



Mobile is assaulting publishing far faster than publishers can cope with the change. It affects everything from search to discovery to publishing platforms to reading. Just as a starting point, mobile demands that publishers revisit their websites to ensure a mobile-friendly design.

Google had a watershed year in 2015. The search engine revealed it received more searches on mobile devices than on desktop computers. It also mandated that “mobile-friendly” pages would get a bump in search ratings, and is also now moving secure sites (https) higher in search rankings.

The future of distribution and access for journal publishers is, inescapably, moving to mobile-first.

Channels

The access options for journals are many and varied. Some authors and journals provide:

- Preprints
- Site licenses
- Single article sale
- Articles in progress
- Bundled access
- Advertising
- Subscription

There is vast discussion these days about Open Access (OA) for journals. Indeed, this is the fastest-growing publication model for new scholarly periodicals. While there are still more subscription model journals (OA journals are about 5 percent of all journals), the OA model penetrates journals that are subscription based since articles within a traditional subscription publication may be OA and make the journal a hybrid. “Born” OA and becoming OA is another consideration: Roughly a **third of articles published** are available free online a year or more after publication.



**Open Access (OA) journals
are about 5% of all journals**



The future of distribution and access for journal publishers is moving to mobile-first.



Chapter 4: Digital Asset Management (DAM) and Archiving

Top Challenge: Journals must work hard to identify formats, structures, and archiving procedures that can span centuries, even millennia.

When we considered the core mission of journal publishing, to record and disseminate new science, we concluded with a step as essential as any other: to archive and preserve the record of science for future generations.

We think of archiving as something after the fact. But it imbues everything we do in journal publishing. The record must be readable by generations to come, with unknowable requirements. The standards that we introduce and enforce in the earliest stages of the publishing workflow pay early benefits in efficiency and versatility. The complexity of assessing the durability of standards drives the enormous effort behind digital archiving.



Image Assets

Images are scanned and optimized. A new generation of software tools (such as [Google Cloud Vision](#)) can recognize objects in photos and automatically tag each image with appropriate metadata. A set of human eyes augments the accuracy and the detail. Suddenly a journal publisher can offer or sell access to its rich legacy, a boon to researchers, and an additional revenue stream for the publisher.

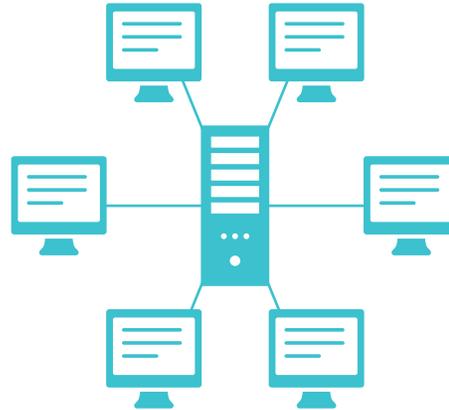


DAMs and Archives

But archiving today is anything but musty. Initiatives like CLOCKSS, JSTOR, and PORTICO recognize that archives require constant refreshing to insure their viability. In this sense archiving and asset management are part of a continuum and are tightly linked. Two simple images illuminate archiving and asset management.

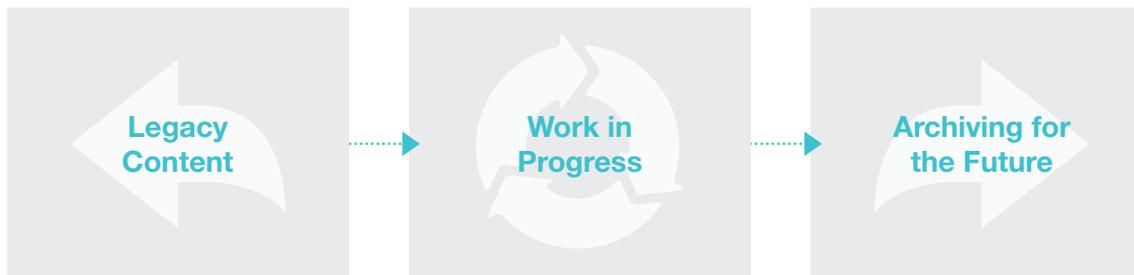
The first you've seen a million times but it still speaks volumes. It represents the essential *centrality* of the network.

It's a star diagram. Regardless of how many devices require access, at the center is the archive, the asset server.



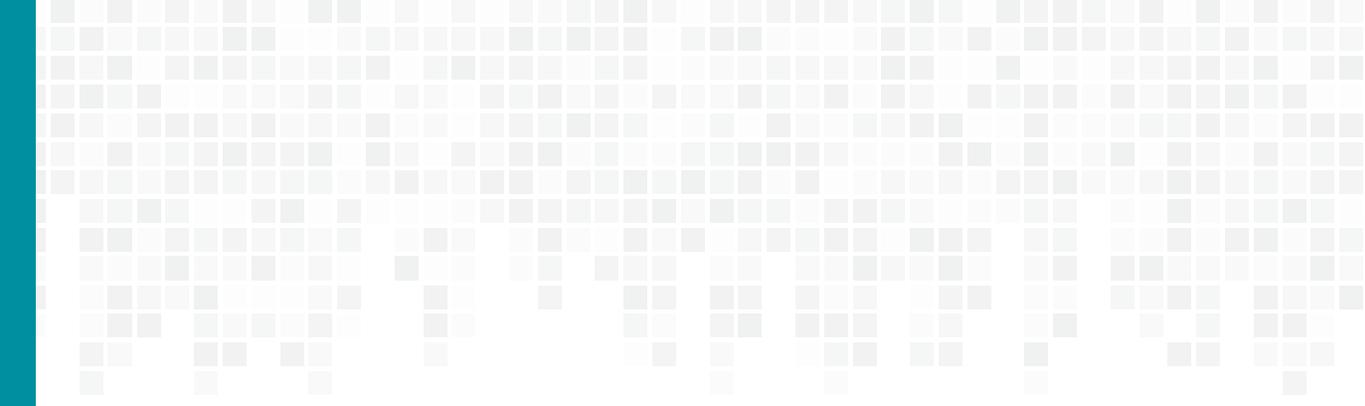
The next image conveys the simplicity of archiving over time, the ongoing nature of the archive.

Managing Archives & Managing Assets



Archiving and asset management are part of a continuum and are tightly linked.





The relationship of archives to the concept of time is a fluid one. Are we pulling material from our dusty archives to present it new and revised? Or, are we pulling last night's unfinished Volume 23, issue 4, from the archive to complete our work today? Just as important is that we file the completed Volume 23 in a secure form so as to be discoverable by future generations of scientists.

A robust digital archive holds the past, the present, and the future. Envisioning the continuum is essential to the successful management of a periodical publication. A monograph is a moment in time. A journal is a living thing, over time.

Mission Critical

Who would you trust to safeguard the legacy of your journal? How will you determine the preservation standards? What is the physical structure of the archive; where are the servers?

Of all of the roles and functions in journal publishing today, archiving shouts loudest for expert assistance. Even with a dedicated archivist on board, journals partner with experts and with a range of digital preservation services including arXiv, bioRxiv, JSTOR, LOCKSS/CLOCKSS, Portico, and PubMed Central.

The future is today.

The Legacy

In accounting terms, legacy assets are a liability. As they age and their relevance deteriorates they become more burden than benefit. Most publishing groups have legacy assets. Sitting inert, in analog form, their value is slight. Some are digital, but in obsolete file formats. Rendered into current digital formats, and tagged with metadata, they come to life. They become a current asset, sometimes of great value.

Legacy assets are mostly text and illustrations. For journal publishers, text predominates. The challenge is to scan the older journal pages, to use optical character recognition (OCR) to digitize the text, to XML tag it, add metadata, and load to an archival server. It sounds like a lot, but many composition vendors have deep expertise with this task.

The American Chemical Society (ACS) digitized nearly half a million original research articles—some 800,000 pages—published between 1879 and 1995. The complete backlist of American Society for Testing and Materials' (ASTM) nine journals have now been digitized back to the first issue of each.



Chapter 5: Identifying Great Partners

Top Issue: Journals face the challenge of choosing the right partner or vendor.

Earlier in this guide we talked about journal publishers “healing” themselves, looking inside to find opportunities for improvement. Organizations often wait until something breaks before they look to improve workflows. Resistance to change is common. Smart organizations are constantly monitoring, knowing that things can always be better.

The business of publishing is becoming increasingly tied to digital technology. But digital isn’t always the answer. It’s essential to differentiate between what can best be automated and what should still be done manually. You don’t have to do everything, end-to-end, before you get started.

Businesses are run by people and they become like people, a collection of features and flaws, strengths and weakness, tics and triumphs. But organizations, like people, often have a hard time looking inside and seeing clearly. We might know that something isn’t working but have trouble identifying the root of the problem. And we’ll be stumped when it comes to solutions. Change moves slowly. It’s time to call for help.



It’s essential to differentiate between what can best be automated and what should still be done manually.





Different Kinds of Partners

There are three situations when reaching out makes the most sense.

In some situations, you might reach out only for advice, for insight. We think there's something wrong, but we don't know what it is. A publishing consultant might come in to evaluate systems and make some gentle recommendations.

Sometimes a publisher identifies a particular piece of hardware or software needed to enhance the flow, new servers perhaps, or an upgrade to the latest version of the Microsoft Office suite. They seek competitive bids from skilled suppliers. Upgrades are installed in running order.

Most importantly, there are vendors supplying services, everything from cloud hosting, ecommerce facilities to full journal production. These vendors are there every day, monitoring, updating, making sure that things run smoothly. And so they are more valuable than one-off suppliers. They become partners. Keeping these firms at arm's length — having them do only what you tell them to — is missing out on much of the value they can provide to your organization. Ideally it becomes a collaborative relationship that allows both partners to be successful.



Keeping partners at arm's length is missing out on a lot of the value they can provide to your organization.



Need a little help?

The Handbook of Journal Publishing identifies different types of vendors/products/services in use by journal publishers. **Some vendors have multiple roles, providing multiple services, in the following areas:**

- Authoring and citation management systems
- Online submission and peer-review systems
- Word to XML text conversion services
- Hosting platforms
- Aggregators
- Abstracting and indexing services
- Subscription agents
- Sales agents
- Distribution houses
- Publishing consultants
- Document delivery suppliers

How Do You Choose?

When searching for a partner, look for a veteran who has experience at the very heart of the journal publishing process. They reach back into authoring and peer review. They facilitate the publishing workflow. Tagging and output is a no-brainer.

Do their skills and experience include:

- Coding and composition
- Copyediting
- Illustration
- Project management
- Editorial subject matter experts
- Custom software creation

Selecting and managing your partnership is an essential function of successful journal publishing today.



When searching for a partner, look for a veteran who has experience at the very heart of the journal publishing process.



Making Friends

Companies are increasingly deliberate in their interactions with suppliers. Oxford University Press **provides a 12-page document** that covers issues like conflicts of interest, protection of intellectual property as well as a range of other services it expects from its suppliers.

When seeking a partner, there are some fundamental considerations:

- **Set goals and objectives:** What problem(s) do you think you're trying to solve?
- **Create a short list of partners:** Research the field. Talk to your colleagues about the vendors they work with. Check conference exhibitor lists.
- **Define a request for proposal (RFP):** Work with consultants to find the right choices for your organization. It doesn't have to be a formal document, but you want to be explicit and specific about your needs.

Here are some of the questions that qualify a publishing services partner:

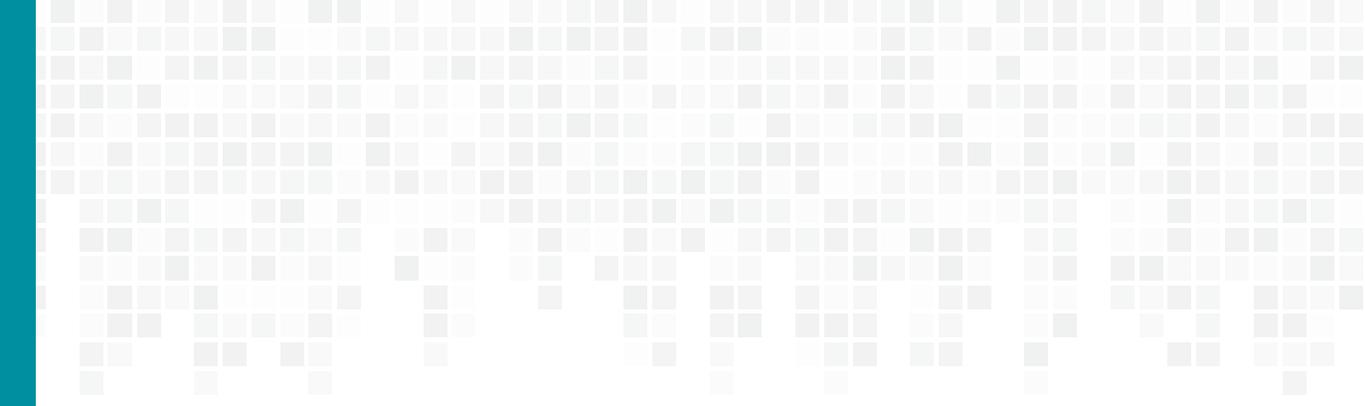
- **Which services can they address?**
- **How long have they been in business?**
- **Who are their key customers? Can they provide solid customer references?**
- **Are journals central to their business, or just also available?**
- **Do they provide unbiased consulting services?**
 - **Are these services billed or included in the editorial and production fees?**
- **How do they manage offshore employees?**
- **How do they manage timely technical support?**

A strong publishing partner will make publishing flourish and can make a difference in your organization.



A strong publishing partner will make publishing flourish and can make a difference in your organization.





The Takeaway

As customers continue to expect articles in new places, different formats, faster, and at lower prices, the publishing industry is doing its best to keep up. They must focus on their workflows.

The word “flow” implies a steady and continuous stream, but publishing workflows are anything but. They’re twisting and choppy. But the concept of flow is essential to managing the publishing process. There are many pieces, but the successful publisher sees links where they exist and forges connections where they don’t.

Publishers that identify their pain points through each stop in the publishing journey—from authoring and editing on through archiving—will discover unexpected opportunities for improvement. They may be able to implement some changes internally. Others will require outside expertise. Whether you must learn new technology or manage changing priorities, there are systems and people who can help smooth the process.



Content & Media Solutions

About Apex CoVantage

Apex is a leading supplier of data conversion, editorial, and content-enhancement services to publishers worldwide. By balancing leading technology and broad industry expertise, Apex delivers exceptional content and media solutions to its customers. Apex has been managing large-scale digitization projects for libraries, publishers, and media companies around the world for nearly 30 years.

 +1.703.709.3000

 content@apexcovantage.com

 @apex_content

 [linkedin.com/company/apex-covantage](https://www.linkedin.com/company/apex-covantage)

 [facebook.com/ApexCoVantage](https://www.facebook.com/ApexCoVantage)

apexcovantage.com

Apex CoVantage
200 Presidents Plaza
198 Van Buren Street
Herndon, VA USA 20170-5338



Interested in learning more about overcoming your publishing challenges?
Contact us today. apexcovantage.com