I T H A K A

I T H A K A R E P O R T
C O M M I S S I O N E D B Y J S T O R

Scholarly Communications in the Biosciences Discipline

A Report Commissioned by JSTOR

Created March 26, 2007
Ithaka Strategic Services
Authors:
Meredith Quinn and Jennifer Kim

NOTE

This is an edited version of a study commissioned by JSTOR in 2006. It is being shared as the research presented here may be of interest to others in our community or beyond. For more information about this report or others in this series, please contact Heidi McGregor, Director, Marketing & Communications (heidi.mcgregor@jstor.org).

I. EXECUTIVE SUMMARY

This study of the biological sciences is the third in a series of discipline reports undertaken by Ithaka on behalf of JSTOR. The purpose of this paper is to explain how research is conducted and disseminated in the biological sciences, and to identify the relative importance of various resources.

The biological sciences are best understood as two clusters of subdisciplines: "bench" subjects and "field" subjects. In our research, we consistently found that bench and field biology differed from each other in meaningful ways, as summarized in this table:

Bench sciences	Field sciences
• Organized around labs • Research problems are defined by the head of the lab • More likely to have biomedical applications • More likely to have funding – especially when research is tied to biomedical applications • Sense that historical journal material (>10 years old) is irrelevant (though it might be cited in article introductions to provide context) • Speed to publication is more important; researchers are afraid of being "scooped" • Higher growth • Higher proportion of commercial publishers	• Organized around labs, but researchers (even graduate students) are more free to define their own research problems • Less likely to have funding • Older journal material (naturalistic descriptions, anatomy) can provide important data for research – but is not widely consulted • Books more likely to play a role, albeit still a small one • Less time pressure to publication • Higher proportion of nonprofit publishers

Across the biosciences, the journal article is the primary medium of communication and the most important criterion for tenure. The scholars we interviewed only share their work pre-publication by sending it to trusted colleagues or by presenting limited findings at conferences. Books are not an important means of communicating research in the life sciences. They have a limited and well-defined place within the field sciences, where they serve an integrative role in teaching graduate students. In the bench sciences, books serve as references, and these are moving online.

Biology articles tend to have several authors. Collaboration takes place in very different ways that our interviewees told us were hard to predict. Some partnerships are very tight; others are quite distant. Sometimes scientists with similar expertise work together; at other times, scientists seek out those with complementary expertise. The dynamic of collaboration within a single lab is different than the collaboration between two labs in two different cities. Collaborations tend to last over a sustained period of time, with many papers written jointly over that time period.

In general, biologists have uses for journal articles at three different points in the research process: while formulating a research question, while conducting research, and while writing. When they search the electronic literature, biologists start with search vehicles that point them to articles, and not with the content sites themselves. Most interviewees reported using Google Scholar or the search feature of Web of Science as a first step. Those who work on medical topics start with PubMed (e.g. an epidemiologist, a developmental toxicologist, and a biophysicist who is housed by a medical school). Vertical searches are also popular, depending on the nature of the question. As a result of this search method, biologists are not very familiar with particular content platforms, including JSTOR. They are indifferent on the question of who actually hosts the articles and they have difficulty recalling the names of content platforms unless they are asked specifically about them.

Once they have found an article, biologists need to assess quickly how relevant it is to them. If an article is helpful, its references are the guideposts for further searches. If an article looks relevant, biologists will download it to their own electronic files or print it out for print files. These personal libraries play a significant part in the research process.

Life sciences librarians acknowledge that going to the library or ordering an article through ILL are truly methods of last resort for biologists. Print resources have almost no place in biological research. To the extent that historical material is valuable, not all is equally so. Based on our interviews, we believe that there are two types of useful historical material in the biological sciences. The first are the "classics" that are necessary for teaching and for writing. These are the must-cite articles that appear over and over again in introductions to papers, to give context to the latest discoveries. The second type of historical article is more relevant for research itself in the field sciences: the naturalistic descriptions that were published in organismal, taxon-specific journals. These are careful, detailed descriptions of a particular animal or ecological system.

We didn't hear a clear consensus on how concentrated the most important journal literature is in the life sciences. Unlike the humanities or social sciences, the sciences have *Nature* and *Science,* two journals in which any of the scientists with whom we spoke would aspire to publish. *PNAS* follows not far behind. The most important life sciences articles make their way to these journals. Beyond that, we heard some indication that ecologists might be more like historians; one ecologist told us that, while there are a few clearly prestigious ecology journals, ultimately he would not be comfortable with just those because he needs to be able to access very obscure material.

II. INTRODUCTION

Ithaka undertook a study of the biological sciences on behalf of JSTOR in January-March 2007. The primary research was based on 26 interviews with scholars, life sciences librarians, and leaders of membership organizations in the field. We also analyzed the relevant findings of the Ithaka 2006 faculty survey and conducted secondary web research.

The report is divided into several sections:

I. Introduction
II. Methodology: describes the resources and methods used for the study
III. Mapping the Field: provides an overview of the structure of the field
IV. Research and Communication: synthesizes findings about how biologists conduct research, share it with each other, and make use of the scholarly literature:
   a. The genesis of a research topic
   b. Collaboration
   c. Communication
   d. When biologists turn to the journal literature
   e. Search and storage
   f. The obsolescence of print
   g. The value of historical material
   h. Concentration of the best work
   i. Non-journal resources
   j. Research in adjacent fields
   k. Frustrations
   l. Open access

III. METHODOLOGY

Our approach consisted of both primary and secondary research. We interviewed 26 people who work in biology, including 19 scholars, five life sciences librarians, and librarians from the National Institutes of Health and the U.S. Forest Service. Our method for choosing interview targets was iterative and based on discipline. In consultation with the JSTOR team, we identified a few disciplines with which to start while we conducted market analysis to identify other promising fields. Based on that analysis, we added more disciplines to the list for exploratory interviews.

Some interviewees were identified in consultation with BioOne. We found the remaining interview targets through web research. We identified the universities that were considered especially strong in our focus disciplines, and reached out to professors on the basis of their biographies on department websites. As a result, our interviewees are located across the United States, including California, Washington, North Carolina, Missouri, New York, Massachusetts, Florida, and Arizona. We sought a mix of professors at various stages in their careers.

Quantitative data from the Ithaka 2006 faculty survey gave context and weight to our qualitative research. For the most part, the two types of data reinforced each other. Where we found discrepancies, we have noted them below.

We supplemented the quantitative and qualitative data with secondary research from a variety of sources, including data sets, websites, and publications. Where appropriate, the sources are referenced in the body of this report.

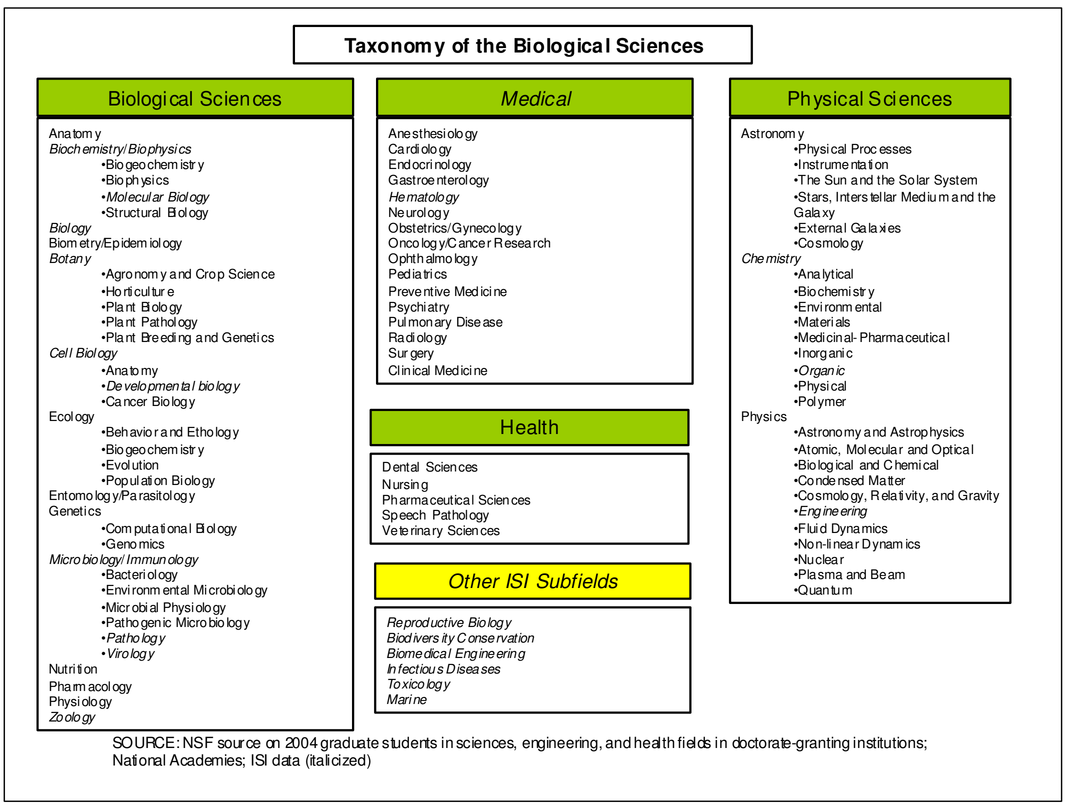
IV. MAPPING THE FIELD

We were surprised to learn that there is no consensus way to describe the biological sciences and the relationship between the subdisciplines. Departments are structured, and scholars classify themselves, in strikingly varied ways. This section describes how, based on our research, we have come to understand the relationships within the life sciences.

Taxonomy

Chart A shows a taxonomy of the life sciences based on the National Science Foundation's (NSF) classification of graduate students. We further adjusted the NSF categories on the basis of the categories used by ISI in its classification of journals.

Chart A



Some notes on Zoology and Botany

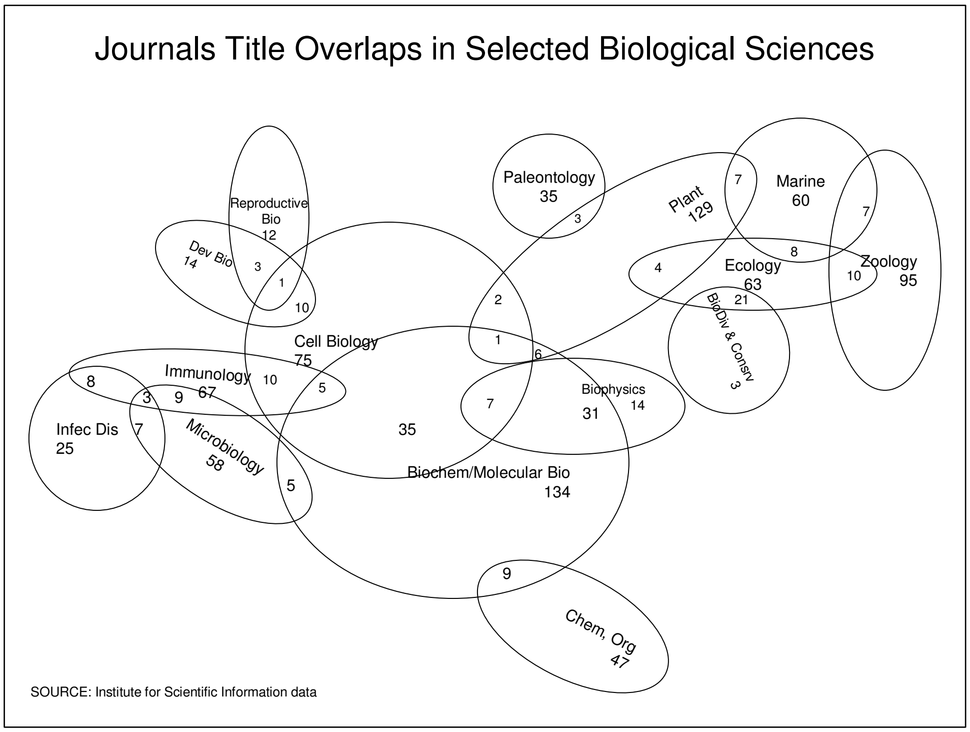
- We heard that "zoology" and "botany" are old-fashioned terms that are increasingly rare on campuses. Many zoology and botany departments have changed their names. For example, at one university, the botany and zoology departments were reorganized into two different departments: the Department of Ecology, Evolution, and Organismal Biology, and the Department of Cellular and Molecular Biology. It made sense to combine botany with zoology because "zoologists" and "botanists" were asking similar kinds of questions, and often questions about the dynamics of plant-animal interaction.
- One librarian told us that the botanists on her campus consider themselves "plant scientists" because "botany" has a connotation of collecting and categorization.

- Zoologists and botanists are more likely to identify themselves by either (1) what they study (e.g. birds), or (2) the kinds of questions they ask.

Overlaps – A tale of two biologies

In order to understand the relationships between life sciences subfields, we examined overlaps in ISI journal categorization. Although most of the journals indexed by ISI are classified into one subdiscipline, a substantial number are classified into at least two categories. We used these classification overlaps to create a rough sketch of relationships between the fields we discuss in this report (Chart B).

Chart B



Note: The total number of titles in any given category can be found by summing all numbers within the circle. For example, there are a total of 112 titles in Zoology (upper right).[1]

With only a few individual exceptions, these relationships were borne out by the interviews and by observation of how departments are structured. Roughly speaking, we can divide the disciplines in Chart B into two areas. The disciplines on the left are the "bench sciences," named for the fact that much of their work is rooted in the laboratory bench setting. Those on the right are the "field sciences," named for the fact that they are traditionally rooted in natural field research. It makes sense that cell biology and molecular biology bridge the two sides, because while they are "bench sciences," their techniques are increasingly being applied in the field sciences as well.

1 For a similar map of all of the sciences, please see http://mapofscience.com/index.cfm?page=1.

The division between bench and field sciences was reinforced consistently throughout our research.[2] The general trends that we observed, which will be described in more detail in the coming pages, are as follows:

Bench sciences	Field sciences
• Organized around labs • Research problems are defined by the head of the lab • More likely to have biomedical applications • More likely to have funding – especially when research is tied to biomedical applications • Sense that historical journal material (>10 years old) is irrelevant (though it might be cited in article introductions to provide context) • Speed to publication is more important; researchers are afraid of being "scooped" • Higher growth • Higher proportion of commercial publishers	• Organized around labs, but researchers (even graduate students) are more free to define their own research problems • Less likely to have funding • Older journal material (naturalistic descriptions, anatomy) can provide important data for research – but is not widely consulted • Books more likely to play a role, albeit still a small one • Less time pressure to publication • Higher proportion of nonprofit publishers

The distinctions between the fields are so pervasive that "life sciences" or "biological sciences" may not be a useful category of analysis. Our impression is that biologists themselves don't think of themselves belonging to a biological science community. One manifestation (and perhaps cause) of this is the hierarchy of journals in which biologists publish. Unlike the humanities and the social sciences, the sciences have flagship publications in which all the best research across the sciences is published – *Science, Nature, PNAS*. The next most prestigious titles in which to publish are at the subdiscipline level – *Evolution, American Naturalist, Ecology, Paleobiology,* etc. With just a few exceptions, there is no flagship, prestigious journal for the life sciences as a whole; there is no equivalent of the *American Historical Review* or the *American Economic Review*. Because JSTOR's biosciences collections are focused in the field sciences, this report places more emphasis in that area.

2 This division is also apparent, for example, in the fact that there are two umbrella organizations for societies in the biological sciences: FASEB (The Federation of American Societies for Experimental Biology, http://www.faseb.org/) and AIBS (The American Institute of Biological Sciences, www.aibs.org). The former serves societies in the bench sciences and those with biomedical applications, and the latter serves those in the field sciences (also called organismal or integrative biology).

V. RESEARCH AND COMMUNICATION

The genesis of a research topic

As in most disciplines, the creative act of designing a research topic stems from the mind of a particular individual, who brings a unique combination of experiences, knowledge and perspectives to his or her field. In one of our early interviews, a post-doc told us that a scientist's research evolved as either new questions applied to the same subject, or the same questions applied to a new subject. This generalization has held up in the course of our interviews. For example, a paleontologist told us that his apparently heterodox research interests in paleontology and astrobiology came about because he was able to apply paleontological approaches usefully to the geology of Jupiter's moons. Conversations with colleagues also play an important role in stimulating new research approaches, especially since biological research is increasingly crossing traditional subdiscipline boundaries.[3]

There are some differences in the extent to which young biologists can direct their own research. In the bench sciences, where research is done as part of a lab, topics are determined by the head of the lab and assigned to graduate students.[4] As the graduate students gain in tenure and eventually become postdoctoral fellows, they might begin to pursue some of their own interests, but still within the scope of the lab where they are based. Research programs in the field sciences might have a similar "lab" organizational structure, but graduate students in those areas have more flexibility to pursue their own interests early on. Professors in these areas function more like advisors do in the social sciences and humanities. This difference presumably comes from the high cost of equipment for the bench sciences, which means that graduate students have to attach themselves to a particular lab and be circumscribed by the interests of the professor whose lab it is.

Collaboration

When you open up a biology journal, no matter what the subdiscipline, chances are that the article you have turned to will have several authors.[5] As one paleontologist put it, "There's no way to do really good science without collaboration. It's the norm." This is true even in fields like ecology that have traditionally had "lone rangers" out in the field. However, collaboration takes place in very different ways that our interviewees told us were hard to predict. Some partnerships are very tight; others are quite distant. Sometimes scientists with similar expertise work together; at other times, scientists seek out those with complementary expertise. The dynamic of collaboration within a single lab is different than the collaboration between two labs in two different cities. Collaborations tend to last over a sustained period of time, with many papers written jointly over that time period. When asked if there was anything that would make collaboration easier, scientists were hard-pressed to come up with ideas. One initially liked the

3 This is evidenced in the next generation of laboratory buildings, which are now designed with open spaces and low barriers "so that researchers would be constantly running into each other," with the goal of cross-fertilization. "A Glance at the Current Issue of Metropolis: The New Generation of Lab Buildings," The Chronicle of Higher Education (27 February 2007, http://chronicle.com/daily/2007/02/2007022701j.htm, accessed February 2007).
4 One article compares bioscience labs to "small family businesses…The [principal investigator] is responsible for choosing research topics, raising money, juggling budgets, and managing postdocs and graduate students." Richard Freeman, Eric Weinstein, Elizabeth Marincola, Janet Rosenblum, and Frank Solomon, "Competition and Careers in Biosciences," Science (New Series, Vol, 294, No. 5550, Dec. 14, 2001) p. 2293.
5 "…For [science] articles published in the EU, for example, the average number of coauthors per article increased from 3.33 to 4.81 between 1988 and 2003, while articles with at least one co-author from a non-EU country accounted for 36% of all articles in 2003, up from 17% on 1988." Mark Ware Consulting, Ltd., for ALPSP, Scientific Publishing in Transition: An Overview of Current Developments (September 2006) 9.

idea of a shared workspace, like an intranet, but then decided that email actually serves his purpose pretty well.

Communication

Biological research is almost exclusively communicated through articles in peer reviewed journals, which are the most important factors – some would say the sole factor – in tenure decisions. The scholars we interviewed only share their work pre-publication by sending it to trusted colleagues or by presenting limited findings at conferences. "I think every scientist has a constellation of people whose opinions he or she values," remarked a marine biologist. "The papers I send them are not for public use. It's understood that you won't talk about it until it appears in print."

Field scientists are traditionally less concerned about being "scooped" than are bench scientists. One interviewee speculated that this might be the case because field scientists are less likely to be researching the exact same topic than are bench scientists. The results of field research are also less likely to be as time-sensitive, which explains why there has not been a push towards speedier publication timelines or publication ahead of print, as we have seen in, for example, genetics journals. We did not find evidence of preprint servers, or a desire to use them, in any of the subdisciplines.

Not all biological research entails the collection of new data sets, but where it does, biologists are particularly concerned with protecting their data. This has hampered efforts to collect datasets into one place where they can be recombined and reanalyzed by others. For example, one professor explained that an ecologist might spend a year collecting the data that was analyzed for an article, and he or she would expect to be able to generate several more articles from the same dataset. At the same time, most subdisciplines have a culture of sharing data on request, so some labs or scientists will make an effort to do as much analysis on a dataset as possible before publishing the first article.

Books are not an important means of communicating research in the life sciences. They have a limited and well-defined place within the field sciences, where they serve an integrative role in teaching graduate students. A paleontologist told us that he thought that even these were declining in value because books don't have impact factors. An ecologist told us that he wished he hadn't written a book early in his career because it did not help him advance towards tenure. The one field science where we did hear that books are valuable in research is in plant biology; "floras" cataloging the plants of a particular region, or monographs categorizing all of the types of, for example, conifers are important tools. In the bench sciences, books serve as reference (e.g., characteristics of a particular material, steps in a protocol), and these are moving online.

When biologists turn to the journal literature

While there are individual quirks, we heard remarkable consistency about the role that the journal literature plays in the research process. In general, biologists have uses for journal articles at three different points of the process:

- While formulating a research question, biologists turn to journals to assess whether anyone has tackled the same question before. They might even write to someone who, on the basis of

previous publications, might be working on the same question, in order to avoid duplication, which was likened to "making a fool of oneself."[6] Uniqueness is very important, and the journal literature serves as a guide to what has already been done. At this point, the searches seem to be more open-ended and exploratory.
- While conducting research (gathering and analyzing data), biologists turn to journals less frequently. They reported doing so when they had a particular question about a protocol or specific statistical technique that another researcher used. One professor referred to this kind of targeted reading as "nitty-gritty analysis."
- While writing, biologists turn to journals in order to write the extensive introductions to their articles. There is a very strong ethos of acknowledging the work of others and placing one's own work within that context. At least a third of the scholars we interviewed used the exact phrase "standing on the shoulders of giants" when describing their approach to the literature. Citations have to be perfect, and for this reason, everyone seems to use a citation management system. Peer reviewers often write back with additional suggestions for citations.

Biologists also keep up with the emerging literature to differing extents and in different ways. Some acknowledged that they did not have time to follow it at all and relied on colleagues or students to give them relevant material. Others scan the table of contents in a hard copy or an e-mail alerting service. The members of a lab might divide up the task of monitoring relevant journals. We did not see any clear patterns in the methods used. This variety is reflected in the 2006 Ithaka faculty study results, in which 33% of biologists said they "often" find information in journals by reading or skimming the important journals in their field; about the same percentage reported that they "occasionally" do this.

Search and storage

When they sit down to search the electronic literature, biologists start with search vehicles that point them to articles, and not with the content sites themselves. Most interviewees reported using Google Scholar or the search feature of Web of Science as a first step.[7] Those who work on medical topics start with PubMed (e.g. an epidemiologist, a developmental toxicologist, and a biophysicist who is housed by a medical school). Vertical searches are also popular, depending on the nature of the question. One paleontologist told us that he turns to GeoRef when his question is about fossil, and to BIOSIS when his question is about modern organisms. "It's handy to have a way of telling them apart." A marine biologist told us that she starts with Aquatic Fisheries Abstracts when she knows that it will have better coverage of the topic she is seeking.

As a result of this search method, biologists are not very familiar with content platforms, including JSTOR. They are indifferent on the question of who actually hosts the articles and they have difficulty recalling the names of content platforms unless they are asked specifically about them. As one scholar put it, "I just hit whatever button seems easiest."

6 We came across two exceptions to this: in epidemiology, it is helpful to have the same experiment run across different populations, for example, an environmental epidemiological study run in five different cities by five different and unconnected teams. In the bench sciences, uniqueness is also very important, but biologists there seem less likely to "cede" topics to other teams.
7 So many said this, in fact, that we were surprised by the 2006 Ithaka faculty study result showing that only 51% "often" or "occasionally" used Google Scholar to find information in journals. On the basis of the interviews, we would have expected this percentage to be higher. We speculate that the percentage would have been much higher were the survey conducted even six months later.

The biologists with whom we spoke avoid keyword searches. Their search process is designed to narrow the number of results that they receive and to increase their relevancy. First, the people with whom we spoke are familiar with the names of the other scientists who are doing research relevant to theirs, and they will often simply search by the names that they know. Some told us that when they know where a particular article was published, they will search for the specific journal on the library's OPAC.

Once they have found an article, biologists need to assess quickly how relevant it is to them. They said that the number of citations to the article or the author was a good first indicator of quality. A few were very emphatic about not valuing a journal's Impact Factor as a signal of quality. The surest way to determine relevance, they said, is simply to read the abstract.

If an article is helpful, its references are the guideposts for further searches. It seems difficult to overstate the importance of reference linking for biologists. In the 2006 Ithaka faculty study, "following citations from other journal articles" was the number one most cited method to find information in journals, and every interviewee said that this was the most important way that he or she navigated through the scholarly literature. They are interested in which articles a given article cites as well as where that article is cited elsewhere. Biologists' reliance on reference linking is enabled by the thoroughness with which each article is expected to cite the sources that provide context for the resource. Where links do not exist in an online database, biologists simply search for the citation in their preferred search engine, though this is "annoying."

If an article looks relevant, biologists will download it to their own electronic files or print it out for print files (in the Ithaka faculty survey, 80% of biologists reported doing this often). These personal libraries play a significant part in the research process. Offline usage is common. In fact, one ecologist and one biophysicist told us that they start their searches within their own reference management systems; another scientist told us that he uses Google Desktop to search his files. (These are more strategies for avoiding excessive results.) We did not hear a historian-like concern with poking into every corner in order to be sure about not missing anything.[8] Offline usage might also be driven by the fact that some biologists do not have continual access to the internet. Ecologists, for example, might spend weeks or months at a time collecting data in the field. The students and postdocs at some labs do not have computers of their own at the lab because the bulk of their work is experimental.

The obsolescence of print

Even life sciences librarians acknowledge that going to the library or ordering an article through ILL are truly methods of last resort. Biologists will first make a calculation as to how helpful the source is likely to be. The scholars with whom we spoke said that it was relatively unusual that they could not find what they needed online, and that when this did happen, it was older articles in more obscure journals that they could not find. As long as one's school has a license, current journals are generally accessible online.

8 One interviewee did tell us that she was afraid she might be missing something when she searches. She wishes that electronic resources made it easier to confirm that she has seen everything that is relevant, but she follows the same reference-liking process as everyone else.

As a result, print resources have almost no place in biological research. Typical was the comment of one zoologist who said, "I don't use print sources – not if I can help it." An ecologist observed that the postdocs at his research center used e-resources exclusively, to the point that, when a snafu prevented their center from accessing issues of *Ecology,* a leading journal, the graduate students simply did not consult it.

Librarians told us that they have been systematically canceling print subscriptions, starting with the biggest commercial publishers. They expressed no nostalgia for print journals. The only time in our interviews that we heard a preference for using print resources was in conversation with the few scholars who told us that they like to browse the current issues of journals in hard copy.[9]

All of this raises the question of why science journals haven't yet eliminated print. None of our interviewees had an answer to this. There was an expectation that this was inevitable – "I can't imagine that any field won't do it" – but they were not sure what the tipping point would be. With the exception of those who liked to browse current physical copies, we did not hear any concern with keeping print titles available. One paleontologist who is involved with two nonprofit publications thought that it all depended upon reaching a certain amount of print cancellations. He was looking to the market to indicate when the time had come. Another paleontologist remarked that it didn't seem to be slated anytime soon. "I've been told for fifteen years that the e-world is just around the corner. That's a very long corner."

One factor that had been slowing the popularity of electronic journals was uncertainty about the quality of digital images, which play an important role in many biosciences. We heard some lingering ambiguity on this point. A paleontologist told us that he thought images were now better viewed online. In contrast, an ecologist told us that digital versions did not yet provide the same detail as print version, and this is particularly necessary in anatomy. Other interviewees saw other benefits to being online, such as the addition of multimedia illustrations to articles, but no one said that online resources in their fields were taking advantage of these capabilities.

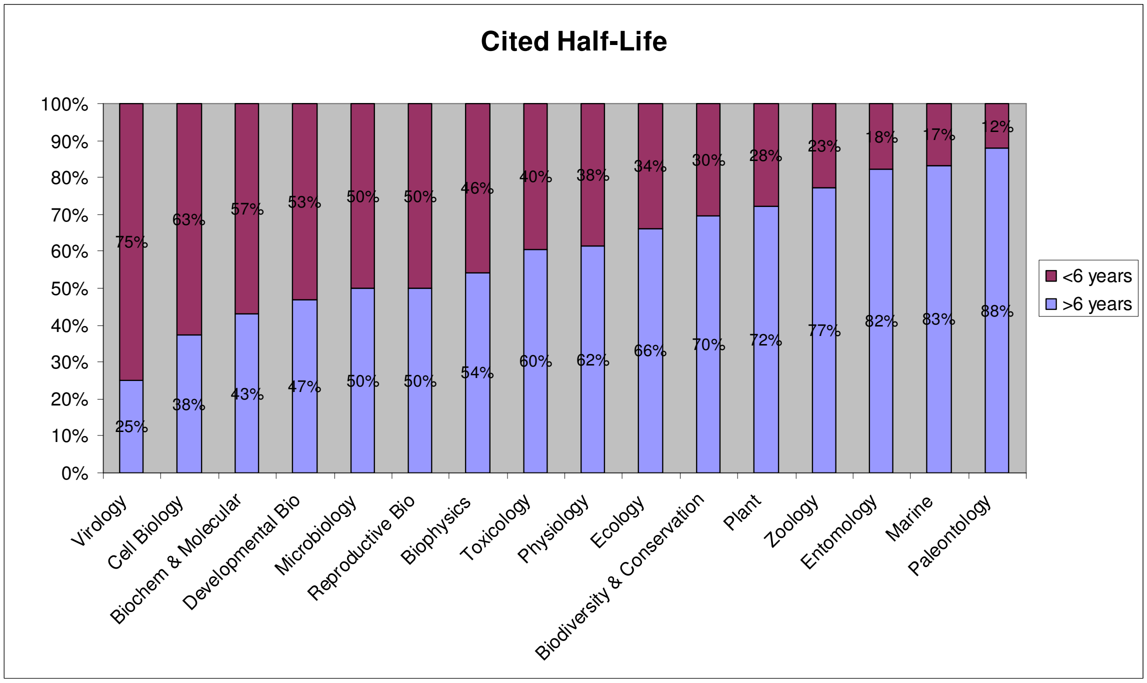
The value of historical material

In the 2006 Ithaka faculty survey, 89% of biologists agreed that "The information in back issues of academic journals is extremely valuable to me." This was exactly the same percentage of historians who agreed with the statement. (Eighty percent of economists agreed.) Based on interviews, however, we believe that way in which back issues are valuable is significantly different for biologists than it is for historians.[10]

As with other questions, the value of historical material varies significantly by subfield of biology, with a clear division emerging between the bench sciences and the field sciences. This is best illustrated by the average cited half-life, as reported by ISI:

9 Most of the scholars with whom we spoke do print out articles that they find online.
10 We also suspect that the definition of "back issue" – in terms of the age – is different for historians than it is for biologists!

Chart C



Source: ISI. This shows the percentage of titles that have cited half-lives that are greater than or less than six years. "Cited half life" is defined as "the number of years, going back from the current year, that account for 50% of the total citations received by the cited journal in the current year."

The bench and medical sciences (on the left in Chart C) simply tend to be faster-moving, and the knowledge those journals contain goes out of date more quickly (particularly since technology in those areas has changed dramatically in recent years). These data were borne out in our interviews, with the most need for currency expressed by epidemiologists and a developmental toxicologist.[11]

Field scientists, on the other hand, generally told us that historical material could be valuable for their research. Paradoxically, most everyone who said this also said that they were more or less alone in their fields in considering historical material important. A zoologist told us that she had been criticized for writing a review article with few post-1950 citations, and ecologists told us that they were forever lecturing their graduate students on the usefulness of historical material, but to no avail. A marine biologist told us that only scholars who are already distinguished can cite older material. Next to the phrase "standing on the shoulders of giants," the most popular phrase we heard in our interviews was "reinventing the wheel" – as in, "Ecologists are constantly reinventing the wheel because we don't know the older material." We heard that this was one area in which JSTOR could have a positive impact, by digitizing material that would otherwise be lost and making it findable. Even as they advocated for more use of historical material, however, field scientists acknowledged that it can be difficult to understand because vocabulary and

11 "For developmental toxicologists, out of 50 references, 48 or 49 will be from the last four years," one interviewee told us. The epidemiologists with whom we spoke also reported that research from the last five years is most relevant for research, although they might cite slightly older work in a grant proposal to explain the history of the research (the trail that got them there).

techniques have changed. There is a certain amount of translation required. On a more practical level, it can also be difficult to keep track of journal titles as they change over time.[12]

Not all historical material is equally valuable. Based on our interviews, we believe that there are two types of useful historical material in the biological sciences. The first are the "classics" that are necessary for teaching and for writing ("whenever possible, you try to hang things on Darwin or one of the old-timers"). These are the must-cite articles that appear over and over again in introductions to papers, to give context to the latest discoveries. A plant biologist, for example, described how the anatomy of a particular plant might have first been defined in 1830, the plant was then studied in various contexts in the 20th century, and today he is piecing together the DNA. His contribution, and the interest for his audience, is in the DNA, but the earlier research provides relevant background. For example, if the DNA suggests particular relationships to other species, he can look back and see if those relationships have been suggested before. Professors also use these as they prepare for undergraduate courses, and they assign them to students in upper-level undergraduate classes.[13]

The second type of historical article is more relevant for research itself in the field sciences: the naturalistic descriptions that were published in organismal, taxon-specific journals (and that are less common today).[14] These are careful, detailed descriptions of a particular animal or ecological system. Some use cases include:

- Then-and-now comparisons. For example, an article from the 1950s might describe a particular wetland area, and could serve as a baseline for a contemporary description. These are particularly valuable in ecology and conservation studies.
- The basic anatomical source for an organism. As one interviewee put it, "The anatomy of a cat doesn't evolve that quickly." Once someone has thoroughly described the anatomy of a frog (as was apparently done in the late 19th century), there is no need to do it again.

In other words, these articles function more as primary sources than as arguments to be built upon or challenged.[15] For comparison, if a historian is doing research on the founding of the United States, she might refer both to arguments advanced in the 1930s and to an edition of the founders' papers published in the 1930s, and either would be acceptable; for biologists, only the latter type – a primary source used as the basis of their own argument – would be acceptable. Beyond saying that taxon-specific journals and *American Naturalist* are more likely to have this descriptive data, however, scholars could not be precise about how to identify which journals are most useful.

Concentration of the best work

We didn't hear a clear consensus on how concentrated the most important journal literature is in the life sciences, i.e., is it more like economics, where a dozen journals cover the core literature,

12 One entomologist cited, in particular, the changing titles of Russian language journals during the Soviet period.
13 In the Ithaka faculty survey, 63% said they "often" use information in e-collections to help them prepare for a lecture or a class; 41% "often" give assignments to students that require them to use electronic collection. This was substantiated by our interviews. One ecology professor told us that perhaps 30-40% of the links he sends his undergrads are to articles found in JSTOR. Graduate education draws upon a wider swathe of the journal literature, as well as upon synthesis texts.
14 For a review of this genre, see "The American Naturalist in American Biology," The American Naturalist (Vol. 100, No. 915, pp.481-492).
15 We heard an exception to this claim from an entomologist, who told us that an article in 1979 had attempted to do a mathematical analysis that was very similar to the one that they were attempting. He found this article accidentally: a colleague in biomathematics, with whom he was collaborating, was visiting another scholar, who told him about the article. This study turned out to be very useful in shaping the substance of their argument.

or is it more like history, where obscure sources can be especially useful? Unlike the humanities or social sciences, the sciences have *Nature* and *Science,* two journals in which any of the scientists with whom we spoke would aspire to publish. *PNAS* follows not far behind. The most important life sciences articles make their way to these journals. Beyond that, we heard some indication that ecologists might be more like historians; one ecologist told us that, while there are a few clearly prestigious ecology journals, ultimately he would not be comfortable with just those because he needs to be able to access very obscure material. (He gave the example of a small journal specializing in California wetlands.) An epidemiologist told us that she could name the 10 journals that she uses most often, but that the epidemiologist down the hall, who has a slightly different specialty, would name 10 journals of which only a few would overlap with hers.

Non-journal resources

Since we heard some variation in the usefulness of non-journal resources, these are presented by discipline below:

Subdiscipline	Other types of resources that are useful, and for what	Comments
Ecology	• Data from experiments – to allow others to verify findings and combine with other sets of data to create meta-analyses	• Extremely difficult to standardize ecological data. ("There are 50 different ways of telling the temperature.") The NCEAS in Santa Barbara has been working on metadata standards. • None of the journals mandate the deposit of data right now, and many ecologists resist it because they plan to do a lot of analysis on one data set and do not want others to be able to do it first. Because the data is often gathered painstakingly in the field, it is proprietary (much like a historian who discovers a trove of documents might not want to make them public until s/he has written a book about them). • It was not clear how having data would enhance the journal articles themselves. In other words, ecologists said they would not need to look at a data set unless the article was very close to the subject on which they were working.

Subdiscipline	Other types of resources that are useful, and for what	Comments
Zoology	• Basic structural databases (e.g. GenBank, the Protein Data Bank) – to provide context to the published articles; to provide comparisons for one's own sequencing • Gray literature – the more applied material, particularly in conservation, appears in governmental and NGO documents. It's not clear how this would be useful for scientific research. The direction seems to go the other way more often: gray literature draws upon the scholarly literature.	• Most journals require that scientists deposit sequences or structures about which they have written into GenBank or the EU's equivalent archive, where it is freely available. • Some resources already link to GenBank and the Protein Data Bank and this can be useful when interpreting the research of others. • A scholar wondered how it would be possible to assess the quality of gray literature.
Paleontology	• Databases of species and fossil occurrences • Gray literature – a source of primary data, for example on geology (from USGS) • Multimedia – in order to better illustrate the ideas behind articles	• Although the databases of species and fossil occurrences were mentioned by paleontologists when we asked about databases that might be useful, none said that they actually used these databases themselves. No one thought it would be useful to link these databases to the journal literature. One interviewee told us that perhaps GenBank should be linked to these databases. • One paleontologist lamented that no online journals had truly taken advantage of the multimedia potential, for example to use sound and animation to illustrate the ideas behind articles.
Epidemiology	• Geographical Information Systems – in order to plot layers of data	• Geographical plotting of this sort has been central to the field since before the advent of computers. It is a core method of analysis and could be usefully linked to journal articles.
Conservation	• Gray literature – a source of practical, applied knowledge; can be more easily (quickly) published than can peer reviewed work	• This is a reflection of the field of conservation, which includes academics, NGOs, and practitioners.

Research in adjacent fields

On the faculty survey, almost 50% of biologists said that they "often" search electronic archives of journals for information outside their areas of expertise. This makes sense given the general pattern of research being a new technique or idea applied to a known area, or a known technique applied to an unknown area. Given our conversations with biologists, however, we would hesitate to call this research "interdisciplinary." It seems to occur primarily in adjacent fields. Our interviewees reported that when they are venturing into a less familiar field, they will often consult colleagues about which resources would be most useful, and then follow the citations forward and backward in the usual way. Some said they run searches on Google Scholar or ISI, but are wary of the avalanche of results that they might get back. One ecologist told us that synthesis review books can be helpful when working in an adjacent field. Finally, crossing into adjacent fields is a frequent motivation for collaboration with other scholars.

Frustrations

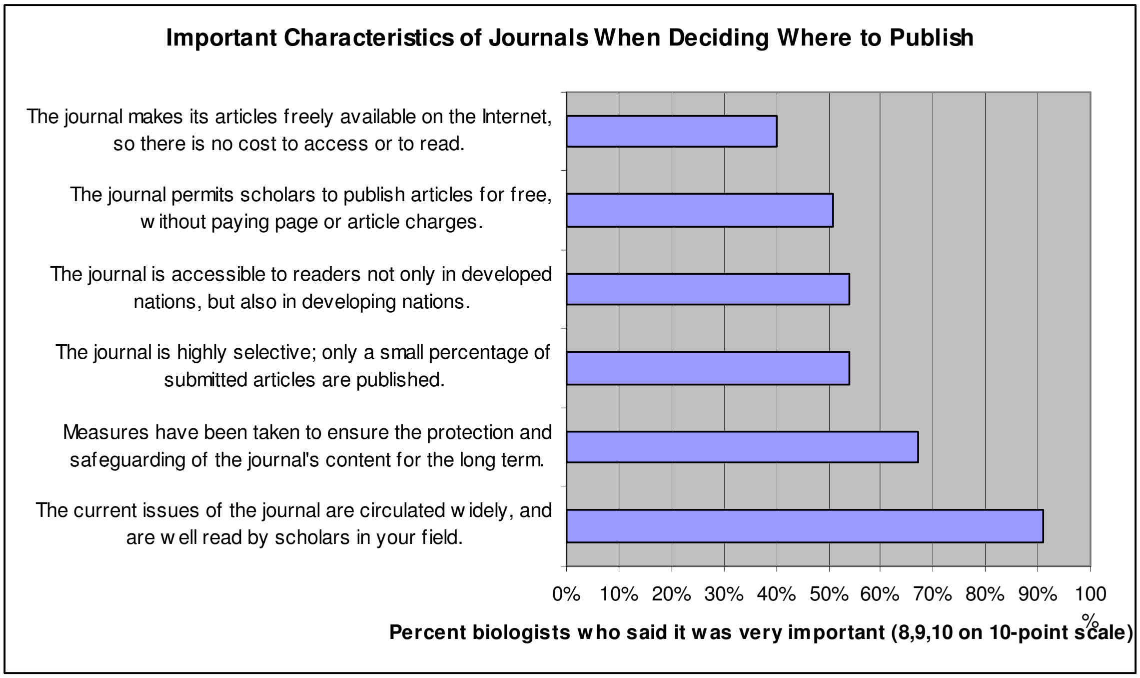
We did not hear a lot of frustration about the knowledge resources available to support research. Interviewees were delighted at how easily they can access resources, and they could not think of ways in which they wish their libraries would support them better. Among the only frustrations we heard were:

- As journals have proliferated, there is "a worry that you're missing something because it's not in the databases you're looking at." It would be useful to see results like: "other papers by this person," or "papers that cited this paper," including results from other journals and other databases (computational biology).
- Not all of the material is online yet. Society journals can be difficult to find online (toxicology). The same is true for more obscure naturalist journals (ecology).
- It is annoying to have to click through many different pages to get to the full text. When a scholar clicks on something from a search engine, s/he wants the full text to come up right away.

Open access

No one with whom we spoke said that the open access (OA) status of a journal had influenced his/her decision about where to publish. This finding echoes the results of the Ithaka faculty survey, in which open access was the least important factor of many when biologists were asked what they considered in their publishing decisions:

Chart D



As was seen across the disciplines in the survey, biologists most highly value that their peers see their work.[16]

Most scholars were aware of open access, and at the very least voiced a vague support for it. The exceptions were two scholars who were very involved in scholarly publishing and believed that "someone's got to pay," and two scholars who are very dedicated to open access and serve as evangelists for it. A few issues that came up:

- The scientists with whom we spoke are not facing a lack of access. "I don't think about [OA] that much, because I have the access I need."
- Even if OA is not the primary determinant of where to publish, the commercial vs. noncommercial status of a journal can be a deciding factor when all else is equal (this was expressed most commonly in paleontology). It was not clear, however, that "all else" is equal very often.
- We heard mixed impressions on the quality of OA journals. A librarian who serves on the tenure committee commented that they are starting to see more publications in OA journals. These assistant professors "don't seem to be concerned about diminished prestige, but they should be." On the other hand, a post-doc told us that *PLoS Biology* is the equivalent of *PNAS* in her mind. According to ISI, *PLoS Biology* had the highest percent increase in total citations in the field of Biology and Biochemistry.[17] However, the faculty study indicated that BioMedCentral is better-known and better-used than is PLoS.[18]
- Prestige still trumps all else: "When I publish, I think about my grad students and postdocs and what's going to help their CVs look better. I'm pragmatic." An ecologist told us that he advises his students to do whatever will have the greatest impact.

16 Still, biologists were more likely to value open access (40%) than were economists (28%) or historians (23%).
17 InCites, November 2006 (http://www.in-cites.com/journals/PLoSBiology.html, accessed March 2007).
18 According to the Ithaka faculty study, 66% of biologists are aware of BioMedCentral; 43% use it. Fifty-three percent are aware of PLoS and 32% use it.

In other words, the average attitude towards OA seems to be “it sounds like a good thing, but it’s not as important as other considerations in where I publish.” We expect that if OA journals were able to establish higher Impact Factors and citation rates *because* they were OA, the career considerations would be a very powerful motivator and could lead to dramatically increased contributions to OA journals.