Using Library Databases to Locate Published Scientific Information

A tremendous amount of scientific information is published every week. As a scientist, you will face the continuing challenge of keeping up with the scientific literature. It's important to recognize early on that available sources of information differ greatly in quality. The most fundamental distinction we make is between information published in refereed journals vs. information published in non-refereed sources. A refereed publication is one that has been reviewed and approved by several experts in the field. Almost all scientific journals publish papers only after several (usually anonymous) reviewers have agreed that the information contained in the paper is both credible and important.

For instance, suppose Dr. Darwin wants to publish some of his recent research on interactions between moths and orchids. After gathering data and writing the paper, Dr. Darwin sends the paper (called a "manuscript") to the journal of an appropriate scientific journal (let's suppose he submits the manuscript to the Journal of Chemical Ecology). After receiving Dr. Darwin's manuscript, the editor of J. Chem. Ecol. sends the paper to several reviewers, who are scientists (usually faculty at other universities throughout the country) who can objectively evaluate Dr. Darwin's manuscript. The reviewers read the manuscript and send their opinions back to the editor.

Each reviewer recommends either,

(a) the paper be accepted for publication without revision (this is extremely rare),
(b) the paper be accepted for publication with minor revision (this is the best recommendation an author can reasonably expect),
(c) the paper be accepted for publication with major revision, or,
(d) the paper be rejected (this is probably the most common recommendation for all journals).

The editor of the journal makes a decision based on the recommendations of the reviewers. Very few manuscripts are accepted without revision the first time they are submitted. Approximately 99% manuscripts are either rejected or recommended for publication only if the author agrees to make the appropriate revisions. Suppose
Dr. Darwin's manuscript is accepted with minor revision. He would then revise the paper to satisfy the comments of the reviewer and resubmit it. The paper would be published approximately 2-6 months later, or even sooner in online journals.

The point of this story is that there is strong quality control on information published in refereed journals (refereed journals are collectively referred to as the "primary literature"). This primary literature can include research articles like that of Dr. Darwin described above and also review articles in major, discipline-specific journals, or in pre-eminent journals such as Nature and Science and journals of the major scientific societies like the British Royal Society and the US National Academy of Sciences.

The two other major sources of published scientific information described below have considerably less quality control. A new category of review article has also emerged in the last 10 years: that of discipline-specific "Trends" or commentary-style articles in journals such as Trends in Ecology and Evolution (TREE). These articles are usually peer-reviewed but they are published quickly with considerable editorial input to ensure that they are up to date with recent developments. These current journals are often a good source of new ideas based on the data-intensive research published in the primary literature.

A second major source of published scientific information is books. Books are useful because they contain a large amount of related information. However, books have two major weaknesses. First, the information published in books does not usually experience the degree of peer review described for journals above (an exception includes widely used introductory textbooks, such as the 4th edition of Begon et al's Ecology, which generally receive considerable scrutiny prior to publication, followed by revisions based on feedback from peers). Second, the most recent edition of a book is often several years old, so recent information may not be included. Please also be aware that the bibliography of the course text is an excellent source of wonderful citations!

The third major source of published scientific information is the Internet. Technically, any information available on the Web is considered "published," though legal issues still remain unresolved. For our purposes, it is important to recognize that most information available on the Web is not peer reviewed, so is not necessarily credible. An exception is information published on the web pages of recognized scientific organizations, such as the National Science Foundation.

An additional source of readily available information is non-refereed publications. These include magazine and newspaper articles, which generally are written by non-scientists and reviewed only by editors who are not scientists. As you might expect, these are relatively poor sources of information that may promote unjustified viewpoints or conclusions and should be considered unreliable. There is a small number of non-refereed, or partially refereed magazines, such as Scientific American, American Scientist, New Scientist and Natural History, that might be
acceptable as cited sources in your term papers. As a rule, however, magazines and newspapers should not be used as sources of scientific information.

The above discussion makes it clear that we should use refereed journal articles (the primary literature) as our major source of scientific information.

There are many ways to track down research articles published in the primary literature. You could start with a textbook, find your subject in the index, read the appropriate section in the book, and look up the scientific articles cited by the author of the textbook. Fortunately a variety of databases have recently been developed to help us locate information in the primary literature. *Biosis Previews, Web of Science, Environmental Sciences and Pollution Management, Scopus* and several databases within *FirstSearch* are all excellent sources of citations. The individual databases in *FirstSearch* most relevant to this class are, *Agricola, Basic Biosis*, and *Biological & Agricultural Science Index*. In today’s lab, we will be guided through the use of these databases, by Carrie Leatherman, your Natural Sciences Librarian (*http://homepages.wmich.edu/~cjm8593*).

**Laboratory Assignment:**

(1) Prepare lists of primary literature articles published in the last 5 years about feeding by larval Lepidoptera (herbivory by butterflies and moths), or choose your own search topic. Your task is to compare the effectiveness of 3 databases at retrieving relevant citations. To do this you should construct a search using key words or subject headings and choose three article indexes. At least one of these should be the library’s “Power Search” and compare its effectiveness with *Biosis Previews* and *Web of Science*. Compare the results you get with the different core databases and try some of the additional databases such as *ScienceDirect*. Do you get the same results if you use the same keywords? Which database do you like more and why? Your comparison should include a table or table plus graphs of numbers of hits and time taken to search for each search term used (this requires thinking about the effectiveness of your search terms).

(2) Once you have compared the effectiveness of the 3 databases you should compile a complete list of at least 10 citations relevant to feeding by larval Lepidoptera, or your own search topic.

(3) Then we would like you to track down three of these articles, either online or in print, and provide annotated summaries of them. The “Find it@WMU” button in the article indexes will direct you to full text online, to the call number of the journal in print or if the library does not have the journal to the library’s interlibrary loan service. Your annotated summaries of each paper in the annotated bibliography should include:
a) the "big picture" that provides the rationale for the study (i.e. why is it interesting?)

b) the major hypotheses under investigation

c) the basic methods used to test those hypotheses

d) the major results, and

e) the relevance of those results to the big picture (i.e. why are the results interesting?).

This summary must be in your own words (don’t plagiarize!) and can include any comments you might want to add about the nature or usefulness of the paper. You should aim at about a third to half of a page for each paper (3 citations = about 1-1.5 pages).

Please hand in as a lab report (40 points):

(1) your 3 searches with key words and a tabulated/graphed comparison of the 3 databases including a conclusion as to which you preferred and why (15 points).

(2) your complete bibliography of at least 10 citations (10 points).

(3) 3 annotated citations (can be part of the total 10) (15 points).