The Limnology
and Oceanography

Bulletin

The American Society of Limnology and Oceanography is a membership-driven scientific society (501(c)(3)) that promotes the interests of limnology (the study of inland waters), oceanography and related aquatic science disciplines by fostering the exchange of information and furthering investigations through research and education. ASLO also strives to link knowledge in the aquatic sciences to the identification and solution of problems generated by human interactions with the environment.

Editors: Adrienne Froelich Sponberg, Ph.D., P.O. Box 8785, Silver Spring, MD, 20907, USA and John Dolan, Microbial Ecology, Laboratoire d'Océanographie de Villefranche, CNRS & Universite, Paris VI, Station Zoologique, B.P. 28, F-06230 Villefranche-Sur-Mer, France

E-mail: bulletin-editors@aslo.org

ASLO Business Manager:
Helen Schneider Lemay, ASLO Business Office, Waco, TX, 76710, USA,
Tel: 254-399-9635 or 800-929-2756,
Fax: 254-776-3767, business@aslo.org
http://www.sgmeet.com/aslo

Advertising: Helen Schneider Lemay, ASLO Business Manager, Tel: 254-399-9635 or 800-929-2756; business@aslo.org

The *L&O Bulletin* is published quarterly by the American Society of Limnology and Oceanography, 5400 Bosque Blvd., Suite 680, Waco, TX, 76710, USA. Postage paid at Waco, Texas. POSTMASTER: Send address changes to ASLO Business Office, 5400 Bosque Blvd., Suite 680, Waco, TX, 76710, USA.

Subscription price to regular members is included in annual dues. Information on institutional subscriptions is available upon request from the ASLO Business Office.

Views expressed in this publication do not necessarily reflect official positions of the American Society of Limnology and Oceanography unless expressly stated.

The *L&O Bulletin* publishes brief articles of broad interest to the ASLO membership, Letters to the *Bulletin* (typically responses to articles), and ASLO News on a quarterly basis. Information on the preparation and submission of articles and letters can be found on the ASLO web site (www.aslo.org). It is recommended that you contact the editors before preparing an article or letter.

THE ENCYCLOPEDIA OF LIFE – NOT JUST ANOTHER WEB ENCYCLOPEDIA!

David J. Patterson, Marine Biological Laboratory, Woods Hole, Massachusetts 02543, USA; dpatterson@mbl.edu

THE BASIC IDEA

From the perspective of a systematist, ecology resembles a multidimensional Mandelbrot game where patterns of order are sought at all levels. This may be why black boxes are familiar within (or containers for) many ecological 'insights'. However, black boxes can hide relationships and effectively prevent a mechanistic understanding of what is going on between players. It is at a reduced level that the 1.8 million known players in the game (the species), and those myriads of unknown ones past and present, determine the transactions and interactions which constitutes the biosphere. The reductionists' model is assembled from these basic elements, tested by and used to test the models that include the black boxes. At this level there is a continuum between ecology and systematics. As pointed out by Charles Godfray (Godfray 2002), it behooves the systematist to find new ways to ease the tasks of the ecologist. Christine Hine (Hine 2008) tells us that the taxonomic community is responding to that call.

To get the reductionist part right, species need to be correctly identified – and not just the sentinels and keystones but the full cast of players. With an inventory of species names, we can, in principle, provide access to all associated knowledge. Until now, there has been no simple way to access this information. Surprisingly enough, we do not even have a place where there is a list of all species. Systematists traditionally have made even this elementary task hard by changing names in response to shifts in systematic and phylogenetic insights over time. Ecologists and others are currently obligated to keep track of these vicissitudes if they wish to demonstrate that they are aware of the prevailing taxonomic concepts, and the correct names for them. The Encyclopedia of Life is a new massive biodiversity project that is set to change things, and among its myriad benefits, can eliminate this chore from our lives.

Since Linnaeus' (1767) achievement of the last major encyclopedic compendium of all life, biology has expanded more rapidly than our capacity to organize the emerging knowledge. Fortunately, the e-world has recently matured sufficiently to provide the means of storage, communication, association, and social collaboration that can culminate in a single portal with a web site for each species (Figure 1) and with access to all information about those species

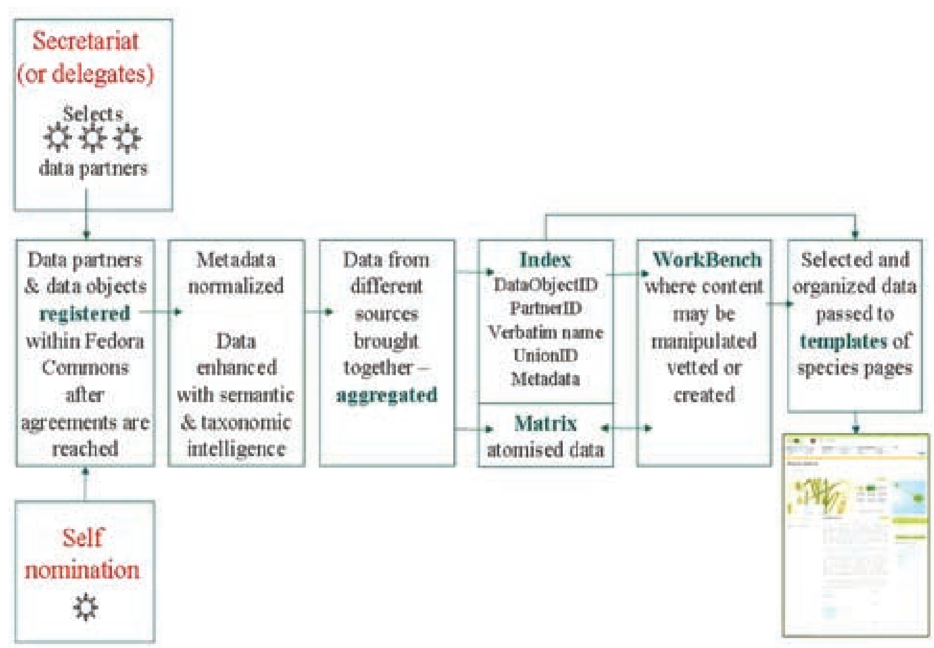


Figure 1. EoL data flow chart.

(Wilson 2003). The Encyclopedia of Life (EOL) project seeks to make this happen. It has made possible through financial support from the John D. and Catherine T. MacArthur Foundation and the Alfred P. Sloan Foundation. It will be both international and collaborative, but its initial course is being charted by the Smithsonian Institution, Marine Biological Laboratory (Woods Hole), Harvard University, the Field Museum in Chicago, and the multi-institutional Biodiversity Heritage Library.

EOL can take advantage of developments not available to previous efforts to realize the same vision. Internet bandwidth has expanded to allow images and other large digital objects to move freely. With the shift to the grid, computing power located remotely can offer synthetic and analytical services – such as Google Earth - that are more powerful and integrative than predecessors that were installed on desktops. Web 2. 0 technologies has had great success in projects such as *Flickr, YouTube, LibraryThing,* and *Wikipedia*. An EOL that is participatory will move quickly and will ensure that the product serves a wider array of users. The capacity to re-use content on the internet for new purposes is illustrated by the results page of a 'Google images' search. This is a composite of information drawn from by scraping from many other web pages. An EOL that uses more advanced 'aggregation' or 'mashup' will not be bottle-necked by the need to re-author information – but can progress quickly by collaborating with thousands of biodiversity web sites and merge pre-authored content. The Dutch national node of the Global Biodiversity Information Facility (NLBIF) data portal (www.nlbif.nl/) among others show that this approach is credible.

WHAT IS DIFFERENT? TAXONOMIC INTELLIGENCE

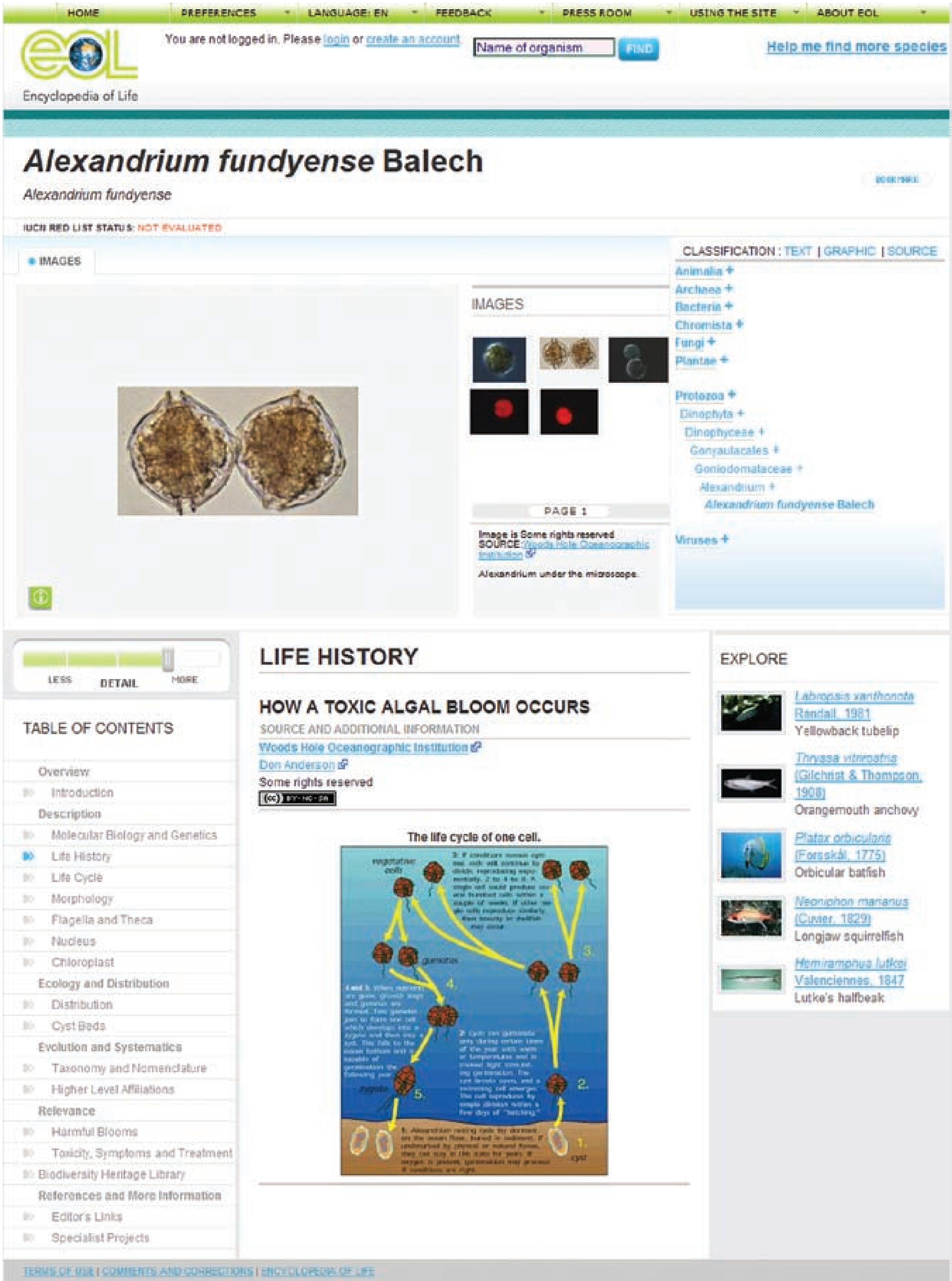
The truly critical advance that makes a unified on-line biology possible has been the intrusion of taxonomic thinking into the design of databases and services. 'Taxonomic Intelligence' refers to a growing collective of strategies that are designed around the expertise and activities of taxonomists. They serve to overcome uniquely biological problems in information management. Names are the key connector between observations and knowledge – in the world of informatics, they are a controlled vocabulary of metadata. Yet, as we all know, the use of names entrains an array of problems – the most significant being that names change with time. As a result, different databases may have different names for the same species. As machines use names to link different pieces of information on the same species together, this problem confounds the assembly of large-scale environment from all of the distributed parts. Until recently, the preferred solution to this names problem has been to push for the adoption of standard names. This is not scalable as we cannot change historical documents to replace out-of-date names with current ones, nor can we sustain the considerable burden of maintaining currency with taxonomic and phylogenetic research at all web sites. Taxonomic Intelligence is a different solution, that of placing all alternative names for the same entity into a 'reconciliation group'.

A query initiated with one name, whether current or obsolete, or whether scientific or vernacular or a mis-spelling or an aberration of machine-driven optical character recognition, can be first passed to the reconciliation group, and then expanded to allow the resulting action to query databases using all available names. If reconciliation is provided as a central service to the Internet, then only that service needs to maintain currency with changing names. The design of this element of taxonomic intelligence is not complex, but the assembly of the system requires over 100,000,000 different names and variants of names to be assigned to fewer than 2,000,000 reconciliation groups – at least. If this were not enough, what constitutes a species is typically not agreed, and so a process and a product must be designed to accommodate diversity and evolution of opinion. This process has begun and will mature over the next 10 years. As it matures, there will no longer be any need for users to keep track of nomenclatural changes, they can type in familiar names and the translation to current name will occur behind the scenes. Other benefits will also emerge. Interested users can receive alerts about new taxonomic concepts (how many entities are there in what we refer to as giraffe), or about new species that have recently been reported in habitats under study. These will promote taxonomic precision and accuracy in ecological research.

Ecologists will be one community who will be able to access biodiversity information that uses this new kind of bioinformatics infrastructure – one that permits a novel amalgam of participation and machine-to-machine communication using web services, one that is taxonomically intelligent, and one that takes us into the domain of the semantic web. The semantic web replaces idiosyncratic home-spun solutions to management of information about organisms on the internet with universal strategies. This will improve the visibility of information by making it more discoverable and so help in the shift from parochial to global. Semantic strategies include the use of persistent and globally unique identifiers for data elements, agreed definitions (metadata), and the common use of protocols to move information around. All of these are now penetrating into biology. Digital Object Identifiers (DOIs are a type of identifier used by the publishing world to refer to books and papers. LSIDs (life science identifiers) are resolvable GUIDs (globally unique identifiers) for biodiversity data objects. Resolvable GUIDs incorporate information that allow us to access the object itself. Services, such as those of CrossRef, can track references from one document to another, and from future documents to past documents. The result is that opening one document unlocks a matrix of cross-linked information – greatly diminishing the workload of discovering information or keeping track of it. Such systems do not need to be limited to documents. Data-sets could be assigned identifiers, and cross referenced – such that one day your efforts to analyse distributed data could be met with messages like: "Other limnologists who used this data set also used the following data sets." Add a little bit of del.icio.us style tagging, and information on quality, relevance, context will also be available.

DOES IT MATTER?

What EOL will deliver will depend on who you are, because EOL is committed to flexibility, offering different composites for different audiences. We can assume that there will be a default 'front page' for a species. At this stage, we expect the front page



Sample EoL page for the *Dinoflagellate Alexandrium fundyense* to show media (images, videos, maybe even 3D zoomable type material), phylogenetic/taxonomic relationships, and text placed in chapters such as: overview; descriptions (which may include an array of descriptions from thumbnails to the original description to suit a diversity of users); ecology and distribution; evolution and systematics; conservation status; relevance, uses (and abuses); as well as additional resources such as web-links, taxonomically indexed RSS feeds from publishers and other media sources (see Figure 1). Content will be presented unadulterated (verbatim) from data partners, although Wiki environments and forum-discussion environments will allow for collective authoring of new versions of text and the formation of derivative versions of content. Content will be freely available. Creative Commons licenses that permit widespread re-use of content (and software) will be favored. It will take at least 10 years for the suite of front pages to be populated, and for the estimated 60% of species that have only ever been observed once, the amount of data may be minimal. Running in parallel and a little behind will be a system that makes deep links into data sets that are accessible through the web and draws selected content out of those resources for display on species pages. Teams within EOL and partnering groups will lead this process.

This is the modest vision. What is driving many of us is knowing that if a names -based infrastructure can build an encyclopedia of all life, then it can index, integrate, and organize any information about any species to contribute to any purpose. Biology can cease to be a parochial discipline of flimsy archipelagos of knowledge stretched thin in time and space. Rather, biology can become a cog in a grand machine. That cog might, for example, be the means of integrating biospheric knowledge with sociological, historical, economic, and geospheric knowledge as we develop full-scale models to explore and challenge global warming.

HOW WILL IT WORK?

A less visible component of the project will be a workbench –a communal and virtual compilation of on-line tools that will allow any user to point to, annotate, associate, visualize, or analyze content (Figure 2). In a sense, the workbench is an open door to participation. The software will be in the form of modules that can interact through appropriate communications protocols and standards. Elements of this environment can be pipelined to allow for the progressive addition of value to data. The software of those modules will be placed in the public domain to allow communal ownership and to give EOL an ever-evolving character.

Early components of the workbench are likely to be places where teams can assemble species pages or sightings environments where folk can record the presence of individuals of any species. EOL will work with other compilers of georeferenced data such as the Ocean Biogeographic Information System (OBIS), the Global Biodiversity Information Facility, GBIF (http://www.gbif.org/), *eBirds* (http://ebird.org/content/ebird/index.html), or *ZipCodeZoo* (http://www.zipcodezoo.com). A simple sightings tool makes it clear that EOL is a participatory environment and opens up the project to widespread participation. Now we can attract data on the distribution and abundance of taxa, especially vulnerable ones, from thousands if not millions of users. Changes to temporal events such as flowering or migrations can be monitored on a scale much grander than was previously possible and correlated with other factors such as global temperatures.

WHAT CAN IT DO FOR YOU?

EOL has the potential to become the device that embraces all biodiversity knowledge, but it will only become that device if it is useful, reliable, and relevant. So what can EOL offer to users? Firstly, that tedious chore of tracking nomenclatural and taxonomic changes can be shed. As a reference work enriched with keys and cross-linked with environmental parameters and georeferences, species identifications will be accelerated, will be more precise, and will be more accurate. Portable high-speed barcoding devices will be able to access EOL and make the bridge between molecular and traditional approaches to biology. If, like GenBank or OBIS, EOL caches information, the resultant growing pool of data can allow research programs to shift from parochial to global, and will facilitate the property of emergence – new science that cannot be predicted from the sum of the parts. To this, add the benefits of the micro-contributions of vast communities of parataxonomists. Having a large pool of data with taxonomic sorting tools would bring back comparative biology that can reveal trends and challenges not even on smaller scales or through the detailed study of a few model species (a macroscope http://radio.weblogs.com/0104369/stories/2002/04/09/macroscope022702.htm). All of this describes a world of research that can be increasingly participatory and collegial – helping a shift from the competitive ethos of the baby boomers. Within ecology, there will be less justification for black boxes, and their number and size will diminish.

WHEN WILL IT BE AVAILABLE FOR USE?

This process became active in February 2008 when the beta version with content on 30,000 species was released for critical input. The remainder of 2008 will be used to gather feedback, appraise priorities, seek data partners, and users. A revised version will be expected early in 2009, and content and infrastructure will continue to grow over the following 5 years.

WHAT CAN YOU DO?

To be part of EOL, there are several things you can do. You can tell EOL developers what you would like to see in EOL (which species, what kind of content, what kind of features and functions), or you can offer content. Content includes text, images, videos, georeferenced data; features include things like – more classifications, comments functions, contributory functions, data visualizations). Please write to me (dpatterson@mbl.edu) in the first instance. Your requests about which species should be included first, the types of content that will be important to you, or the features you would like to see will impact our priorities. Similarly, if you have content ready, and are willing to allow the content to be available under an appropriately open Creative Commons license, then we can make your content visible within EOL. The most important first step in a names-based infrastructure is to establish the names of all relevant taxa within the indexing system, assembling the reconciliation groups, and getting the classification in line with current thinking. The tools for doing this will be available any day now.

REFERENCES

Godfray, C. H. J. 2002. Challenges for taxonomy. Nature 417:17–19.

Hine, C. 2008. Systematics as Cyberscience: Computers, change, and continuity in Science. MIT Press.

Linnaeus, C. 1767. Systema naturae per regna tria naturae, secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis. 13th Edition, I. Thomae Trattnern.

Wilson, E. O. 2003. The encyclopedia of Life. Trends Ecol. Evol. 18:77-80.

THE ETHICS FORUM: THE ETHICS OF CONDUCTING SCIENCE ABROAD

ASLO Professional Ethics Committee; ***Ilana Berman-Frank****, Mina & Everard Goodman Faculty of Life Sciences, Bar Ilan University, Ramat Gan, 52900, Israel; irfrank@mail.biu.ac.il*

[This column is intended to highlight a few issues related to conducting science and field science abroad and is not a comprehensive study.]

As aquatic scientists, many of us have the opportunity to conduct experimental research abroad. This may be at a collaborator's laboratory, at state-of-the-art institutes where regulations are well outlined and kept under check, or at remote field stations, or at laboratories spanning the globe from unpopulated environments to laboratories in developing nations.

The reasons for undertaking research in other countries are varied. Scientists are driven by the scientific quest and exploration of new habitats/organisms or areas under different environmental conditions, and the application of theoretical and laboratory experiments in natural field settings available only in other countries. Scientists also choose their research location by more pragmatic considerations such as access to funding and international collaboration, personal and commercial gains, and in some cases the ability to undertake research that would not be ethically or legally permitted in their home countries (Skene 2007). Whatever the reason, scientists are on the move, with research abroad making up a large part of many grant proposals and scientists' time-tables.

Imagine a seemingly simple scenario that raises a *mélange* of ethically related questions you may encounter on your next research foray abroad. You have a funded project that takes you to a marine station on an island somewhere in the middle of the Pacific. This is an area of great biodiversity and pristine waters, yet it is also the home of and source of livelihood for local populations including scientists that specialize and subsist on local research topics. Your funded grant overlaps with some of these local research projects and may impact populations such as subsistence fishermen or tour operators.

What obligations do you as a visiting researcher have towards the environment and the people? Is your science project designed only to further science, or is it beneficial mostly for your personal advancement? Does it take into account the benefits and/or damages the project will have on the local populations and environment? Does your research comply with local laws? Have you obtained the permits required by national and local authorities for collecting specimens? If experimental manipulation is involved how will it affect this environment on short-and long-term bases? Are local populations and local research activities affected by your plans?

Moreover, are you coming as a collaborator sharing equipment, data, and intellectual property rights? Are you paying for your use of the infrastructure, labor, and equipment? Is authorship on forthcoming publications to be shared with local collaborators, or are they left "to pick up the pieces" while the high-impact publications go to the well-funded scientists?

These questions and other ethically related dilemmas, arising from the situation of the scientist working away from his or her home turf, may be categorized into two broad categories which highlighted here: 1) ethics related to the environment itself, and to our exploitation (utilization) of it to advance science; 2) ethics related to the human environment and populations in the area which we come to discover/explore.

Ethical dilemmas arising from scientific exploration of the environment fall into the realm of environmental ethics. The discipline of environmental ethics has developed since the 1970's, and a large volume of literature is available on all aspects of environmental ethics. While environmental ethics is traditionally a "land ethics" (Dallmeyer 2003), human investigations and activities in aquatic systems, such as whaling and fishing, warrant the expansion of these traditional themes (Dallmeyer 2003). Moreover, despite the boom in environmental ethics, there are still apparently very few ethical guidelines available for ecologists and field (aquatic) researchers (Farnsworth and Rosovsky 1993; Marsh and Kenchington 2004). This is relevant for research both at home and abroad. In their review "The role of ethics in experimental marine biology and ecology", Marsh and Kenchington (2004) argue that, these research areas lack definitions, for the most-part, on what constitutes appropriate (ethical) behavior for scientific exploration in areas such as biodiversity, habitat integrity, community dynamics, and manipulative schemes such as Fe-fertilization experiments. *"As a result, most experimental marine biologists and ecologists operate without ethical guidelines or scrutiny, despite intermittent community concern about their activities in response to specific controversies… This contrasts with the abundant and strict ethical guidelines available in many countries for health related research as well as that involving humans subjects and other sentient animals."* (Marsh and Kenchington 2004).

Thus, while at home legal bounds may protect and delineate what is permitted in terms of environmental and field research, the lack of ethical guidelines, combined with loose, or the absence of, legal regulation in many countries, may foster unethical scientific behavior. Scientists going overseas may collect rare specimens, apply manipulative treatments (changing pH, adding nutrients, radioactive materials, herbicides, etc.) and harm sensitive environments, and in general assume that their activities are justified for the sake of science. Marsh and Kenchington (2004) point out evolving measures that intend to limit unethical science and encourage appropriate behavior in experimental field work. These include ethical standards required from authors