

Editing the Encyclopedia of Life

Misha Kidambi

"Imagine an electronic page for each species of organism on Earth, available everywhere by single access on command," wrote Edward O Wilson (professor emeritus, Harvard University) in an article published in *Trends in Ecology and Evolution* in 2003. His grand vision is being put into practice, and information on every known species on our planet will find its way onto the giant virtual platform of the Encyclopedia of Life. The first version (www.eol.org) went live on 26 February 2008. The entire project will require about 10 years for completion because, apart from developing a site for each of the 1.8 million species on which complete taxonomic information is available, the Encyclopedia of Life portal will have provisions to include information on the millions of species that are continually being discovered. Ensuring and maintaining the quality of all these pages will be quite an editorial challenge.

A Mammoth Project

The two dozen exemplar pages—including pages on the death cap mushroom (*Amanita phalloides*), the tomato (*Solanum lycopersicum*), the imperial blue butterfly (*Jalmenus evagoras*), and the peregrine falcon (*Falco peregrinus*)—give a peek into what the completed pages in the encyclopedia look like. The entry page has links that will take users to pages with information about the organism's evolutionary history, taxonomy, physiology, molecular biology, ecology, diseases, and life cycles, including images, charts, and even videos. All this adds up to quite a large number of electronic pages!

MISHA KIDAMBI, a graduate student in science and technology journalism at Texas A&M University, prepared this article while a Science Editor intern.

A statement made by Gerald Weissmann, the editor-in-chief of the *FASEB Journal*, gives an idea of the vastness of the project: "It is expected to fill about 300 million pages, which, if lined up end to end, would be more than 83,000 kilometers long, able to stretch twice around the world at the equator." And given Wilson's vision that each page will be infinitely expandable and its contents continuously peer reviewed and updated with new infor-

"constantly evolving encyclopedia that lives on the Internet, with contributions from scientists and amateurs alike"

mation, one might wonder what goes on behind the scenes to gather information for this mammoth project and ensure its quality and consistency.

Gathering Information

Using what is known as mash-up technology, the Encyclopedia of Life has brought together available information about species from various providers. Says Breen Byrnes, public information officer for the Encyclopedia of Life project, "the scientific and common names of a species may come from one data provider, photos from another, maps from a third, links to genetic information about the species from a fourth, digitized literature from yet another site, etc." One aim of the Encyclopedia of Life is to be a one-stop shop for anyone looking to find information on any species present on Earth.

David Patterson, the director of the encyclopedia's Biodiversity Informatics Group, says he envisions the encyclopedia's role as much more than being just an "electronic book". In line with

Wilson's vision of creating a "constantly evolving encyclopedia that lives on the Internet, with contributions from scientists and amateurs alike", Patterson wants the encyclopedia to be "a meeting place where enthusiasts can chat with each other, and through this the more knowledgeable can inform the less knowledgeable, and different perceptions and opinions can be aired and perhaps resolved".

Ensuring Quality

With a vision so grand, it becomes imperative to ensure the quality of the information that can be retrieved. Achieving the Encyclopedia of Life vision would mean actively engaging as many people as possible as providers of information, syntheses, critiques, and software, so ensuring accuracy is expected to be complicated.

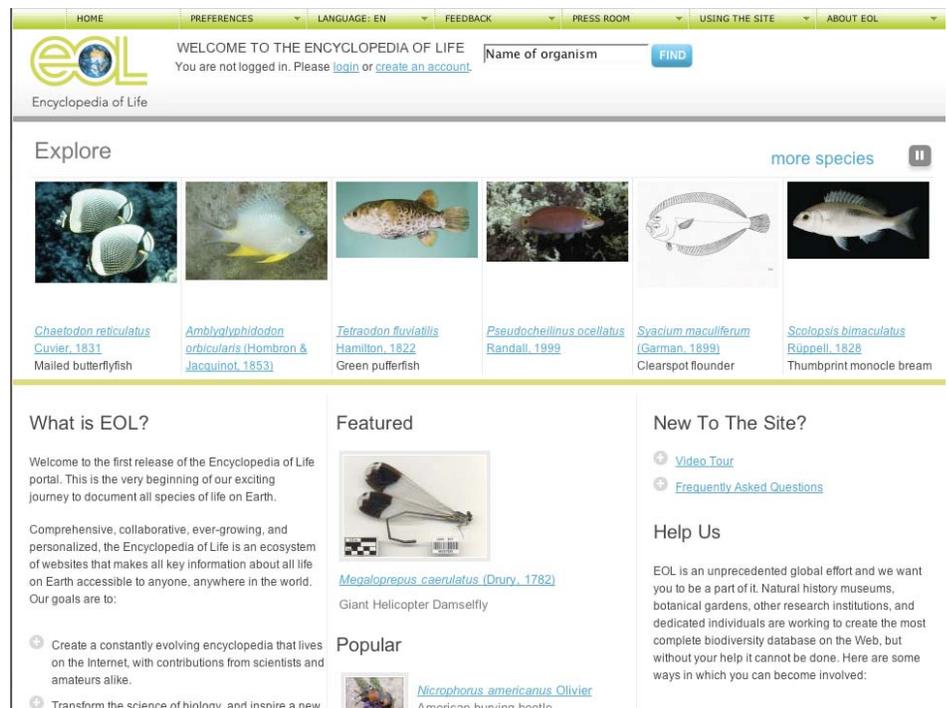
The plan is to achieve quality by more than one means. First, authoritative online Web sites that represent the understanding of the expert community have been chosen as the early data providers; these include such sites as FishBase, Tree of Life, and AmphibiaWeb. Second, readers may use the comments and corrections tab that is provided at the end of every page to comment on any aspect of the content in the encyclopedia, and comments will be passed to custodians who act as editors. "Custodians will include professional scientists, societies and organizations, and amateur enthusiasts who have proved themselves through the quality of prior contributions," Patterson says. Finally, while being worked on, each contribution is marked by one of three quality flags. The first flag is used to indicate data for which quality-checking has not yet begun; the second, to indicate that content is being checked by the custodians; and the third, to indicate that one of the custodians has verified the content.

The custodians work on the contribu-

Encyclopedia continued

uted content and review the comments in a largely online communal environment known as the WorkBench, which permits annotation, commenting, and communal or peer-to-peer discussion. The discussion continues until the custodians are convinced that the content meets a satisfactory standard. Content will finally be visible on the species page once it has undergone the rigorous vetting process. “The process [is] very much the same that has worked very well for the scientific journals,” Patterson says.

When Denis Diderot, one of the greatest encyclopedists, published *Encyclopédie* in 1751, he wanted to ensure that the work of the past would be useful in the future and that the future generations would be better educated for the effort. By combining the power of the Internet with the authoritativeness of an encyclopedia, Wilson envisioned that the Encyclopedia of Life would achieve something as noble: “To transform the science of biology, and inspire a new generation of scientists, by aggregating all known data about every living species. And ultimately, to increase our collective understanding of life on Earth, and safeguard the richest possible spectrum of biodiversity.” 



The screenshot shows the EOL website interface. At the top, there is a navigation bar with links for HOME, PREFERENCES, LANGUAGE: EN, FEEDBACK, PRESS ROOM, USING THE SITE, and ABOUT EOL. Below this is the EOL logo and a search bar with the text "WELCOME TO THE ENCYCLOPEDIA OF LIFE" and "You are not logged in. Please [login](#) or [create an account](#)." The main content area is titled "Explore" and features a grid of six species cards, each with an image and a caption: *Chaetodon reticulatus* (Mailed butterflyfish), *Amblyglyphidodon orbicularis* (Hombrom & Jacquinet, 1853), *Tetraodon fluviatilis* (Hamilton, 1822) (Green pufferfish), *Pseudocheilinus ocellatus* (Randall, 1999), *Syacium maculiferum* (Garman, 1899) (Clearspot flounder), and *Scolopsis bimaculatus* (Rüppell, 1828) (Thumbprint monocle bream). Below the grid, there are three columns of content: "What is EOL?" (Welcome to the first release of the Encyclopedia of Life portal...), "Featured" (Megaloprepus caenulatus (Drury, 1782) Giant Helicopter Damselfly), and "New To The Site?" (Video Tour, Frequently Asked Questions). There is also a "Help Us" section and a "Popular" section (Nicrophorus americanus Olivier American burying beetle).