The encyclopedia of life

Edward O. Wilson

Museum of Comparative Zoology, Harvard University, 26 Oxford Street, Cambridge, MA 02138-2902, USA

Comparative biology, crossing the digital divide, has begun a still largely unheralded revolution: the exploration and analysis of biodiversity at a vastly accelerated pace. Its momentum will return systematics from its long sojourn at the margin and back into the mainstream of science. Its principal achievement will be a single-portal electronic encyclopedia of life.

Imagine an electronic page for each species of organism on Earth, available everywhere by single access on command. The page contains the scientific name of the species, a pictorial or genomic presentation of the primary type specimen on which its name is based, and a summary of its diagnostic traits. The page opens out directly or by linking to other data bases, such as ARKive, Ecoport, GenBank and MORPHOBANK. It comprises a summary of everything known about the species' genome, proteome, geographical distribution, phylogenetic position, habitat, ecological relationships and, not least, its practical importance for humanity.

The page is indefinitely expansible. Its contents are continuously peer reviewed and updated with new information. All the pages together form an encyclopedia, the content of which is the totality of comparative biology.

The rationale

There are compelling reasons to build such an all-species encyclopedia. Not least is the heuristic power for biology as a whole. As the census of species on Earth comes ever closer to completion, and as their individual pages fill out to address all levels of biological organization from gene to ecosystem, new classes of phenomena will come to light at an accelerating rate. Their importance cannot be imagined from our present meagre knowledge about the biosphere and the species comprising it. Who can guess what the mycoplasmas, collembolans, tardigrades and other diverse and still largely unknown groups will teach us? As the species coverage grows, gaps in our biological knowledge will stand out like blank spaces on maps. They will become destinations toward which researchers will gravitate.

For the first time, the biotas of entire ecosystems can be censused in full. Unknown microorganisms and the smallest invertebrates, which still comprise most species yet lack even a name, will be revealed. Only with such encyclopedic knowledge can ecology mature as a science and acquire predictive power species by species, and from those, ecosystem by ecosystem.

As one result, the human impact on the living environment could be assessed in far more reliable detail than is now possible. Today, for example, we base estimates of species extinction on data from a scattering of taxonomically best known groups, including the flowering plants, land and freshwater vertebrates, and a few invertebrates, such as butterflies and mollusks. These taxa contain only about a quarter of the known species on Earth, and almost certainly a much smaller fraction of those still unknown. Tomorrow, other invertebrates, including insects and nematodes, as well as fungi and nearly all microorganisms, together comprising most species on Earth, as well as essential pathways of the energy and materials cycles, can also be assessed.

The all-species encyclopedia will serve human welfare in more immediately practical ways. The discovery of wild plant species adaptable for agriculture, new genes for enhancement of crop productivity, and new classes of pharmaceuticals can be accelerated. The outbreak of pathogens and harmful plant and animal invasives will be better anticipated and halted. Never again, with fuller knowledge of such extent, need we overlook so many golden opportunities in the living world around us, or be so often surprised by the sudden appearance of destructive aliens that spring from it.

An all-species encyclopedia of life is logically inevitable if for no other reason that the consolidation of biological knowledge is urgently overdue. In its earliest stages, already emerging, it forms a matrix within which comparative studies are rapidly organized. The process will accelerate as traditional taxonomic procedures, still mostly dependent on repeated examinations of type specimens and print literature, are replaced by high-resolution digital photography, nucleic acid sequencing and internet publication. With further documentation organized into the species pages, new lines of research will open at a quickening pace. Model species for laboratory and field research can be more easily found -- obedient to the principle that for every problem in biology, there exists a species ideal for its solution.

A growing, single-access species-structured encyclopedia will ease navigation through the immense biological data bases. Aided by computer search engines, patterns can be summoned whose detection would otherwise demand impracticable amounts of effort and time. Principles and theory can be built, deconstructed and rebuilt with an unprecedented power and transparency.

Ultimately, and at a deeper level, the all-species encyclopedia will, I believe, transform the very nature of biology, because biology is primarily a descriptive science. Although it depends upon a solid base of physics and chemistry for its functional explanations, and the theory of natural selection for its evolutionary explanations, it is

Corresponding author: Edward O. Wilson (ewilson@oeb.harvard.edu).

http://tree.trends.com 0169-5347/02/$ - see front matter PII: S0169-5347(02)00040-X