

New Species

Top Ten 2007

Most people do not realise just how incomplete our knowledge of Earth's species is. In 2006, for example, an average of nearly 50 species per day were discovered and named, totalling 17,500 new species!

To attract people's attention to the discovery of species, an international committee of experts has just published a list of the 10 most curious and surprising species described in 2007. The committee, chaired by Dr. Janine Caira of the University of Connecticut selected the Top 10.

Electrolux addisoni, an electric ray, is ranking first. The name is evidence that taxonomy doesn't lack a sense of humour. It alludes to the ray's impressive display of light and to its mode of feeding. Its vigorous sucking action rivals a well-known electrical device used to suck the detritus from carpets in modern homes.



Electrolux addisoni



Gryposaurus monumentensis



Desmoxytes purpuresea

Second best is ***Gryposaurus monumentensis***. This new dinosaur species is one of the most robust duck-billed dinosaurs ever found, and probably the largest dinosaur in the 75 million-year-old Kaiparowits Formation fossil ecosystem.

A shocking pink dragon millipede, ***Desmoxytes purpuresea***, from Thailand is at number 3.

And a mushroom made it onto the list, perhaps because it was discovered in one of the most intensely studied floras in the world, on a campus belonging to Imperial College, London. ***Xerocomus silwoodensis*** illustrates how poorly species are known.

Not to forget plants, the taxonomists who chose the Top Ten put ***Tecticornia bibenda*** at place 10. This plant, found in Western Australia, looks inflated like a Michelin Man. It's one of 298 Western Australian plant species described in 2007.

-KHO-

Protein mining

Million Minds

WikiProteins, a new wiki-based collaborative approach to community annotation and knowledge sharing, was launched on May 23rd by the WikiProfessional initiative. The product of a collaboration including Amos Bairoch of UniProt, Michael Ashburner of the Genome Ontology Organisation and Jimmy Wales, the co-founder of Wikipedia, the aim of **WikiProteins** is to combine the power of text mining engines with community-based annotation and curation.

WikiProteins scans existing databases, such as PubMed and SwissProt, and ex-

tracts concepts such as proteins, processes and pathways using state of the art text mining engines. The relationships between these concepts are then stored in "knowlets", the basis of the **WikiProteins** system. Users can explore the network of related knowlets, thereby getting a continuously up-to-date picture of our current state of knowledge about a particular topic.

WikiProteins is based on the wiki approach: entries are open to editing by professionals and members of the public, although material from the original sources is protected from alteration. This allows **WikiProteins** to overcome some of the limitations of automated data mining.

Text mining is an increasingly powerful tool and is becoming ever more important to scientists facing the explosion of publications. But text mining faces serious limitations. For instance, mining databases for the yeast protein CLB2 (the gene for G2/mitotic-specific cyclin-2) returns not only the correct entry, but also a synonym for emb-9, which encodes the collagen alpha-1(IV) chain in *Caenorhabditis elegans*, and even a dental etching compound. By having the scientific community curate and add to the results of WikiProfessional's automated mining, it is hoped that a "million minds" approach to data annotation and curation will lead to aggregated expert knowl- ▶▶

BY RAFAEL FLORÉS

PAUL THE POSTDOC

