

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-

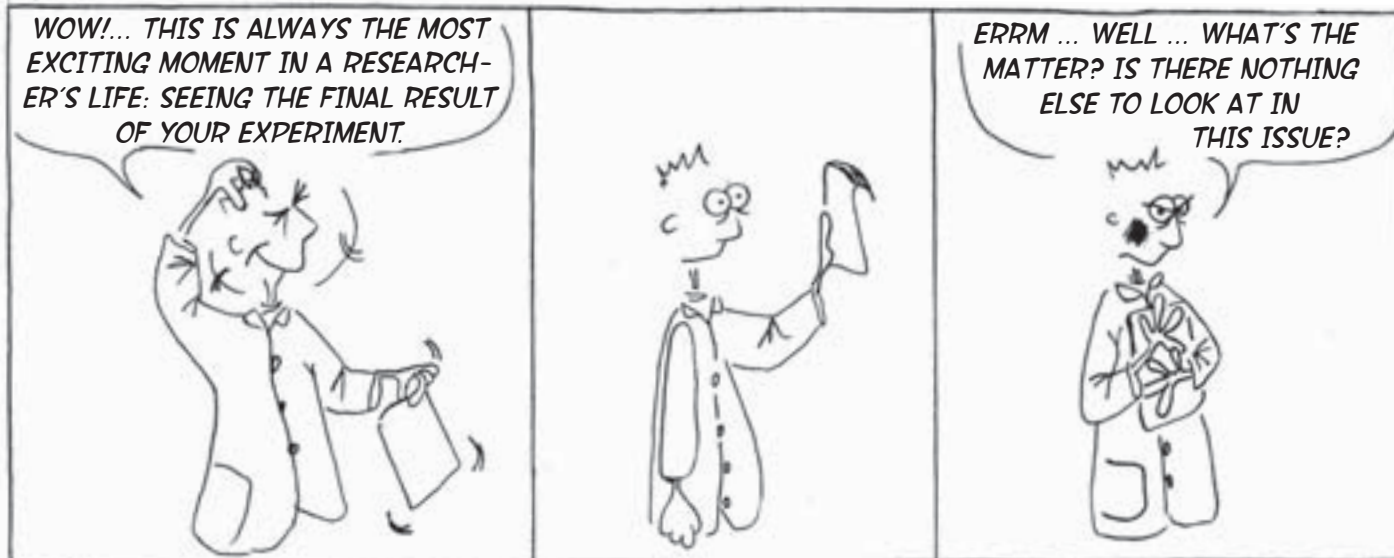
tracts concepts such as proteins, processes and pathways using state of the art text mining engines. The relationships between these concepts are then stored in "knowlets", the basis of the **WikiProteins** system. Users can explore the network of related knowlets, thereby getting a continuously up-to-date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



Recently Awarded

► **Maria Blasco**, Leader of the Telomeres and Telomerase Group at the National Cancer Research Centre (CNIO) in Madrid, will be presented the Körber European Science Award on 8 September in Hamburg. In 1997, while doing a post-doc in Carol Greider's lab at Cold Spring Harbor Laboratory, she cloned one of the mammalian telomerase genes and generated the first telomerase knockout mouse. Since then she has contributed significant insights to the concept that telomerase activity plays decisive roles in ageing and cancer susceptibility.

► **John O'Keefe**, Professor of Cognitive Neuroscience at University College London, will receive the 2008 **Neuroscience Prize of the Peter and Patricia Gruber Foundation**, worth \$500,000. In 1971 O'Keefe, together with colleague Jonathan Dostrovsky, first described the existence of specific neurons in the hippocampus that play a major role in spatial navigation. He suggested that these place cells might form the basis of a cognitive map of the environment. In subsequent experiments, he and others were able to confirm this hypothesis, particularly by demonstrating the collective and cooperative behaviour of the place cells in order to identify the environment.

► The Royal Netherlands Academy of Arts and Sciences has announced this year's laureates for the **Heineken Prizes**. The Prize for **Biochemistry and Biophysics** will be presented to **Jack Szostak** from Harvard Medical School in Boston. Szostak has been involved in many achievements, among them, for example, the construction of artificial chromosomes in yeast. More recently, he and his team successfully evolved a new enzyme that previously hadn't occurred in nature.

The Prize for **Cognitive Science** is being awarded to **Stanislas Dehaene**, director of the Cognitive Neuro-imaging Unit at INSERM in Paris. He has significantly contributed to the understanding of the mechanisms by which our human brain copes with mathematical and semantic challenges.

► edge that cannot be gained by database searches or text mining alone. In addition, researchers can enter sentences from full text articles for the *WikiProteins* engine to incorporate into the knowledge cloud. The *WikiProteins* team estimates that about half of the protein-protein interactions are hidden behind the "firewall" of copyrighted journals.

Three principal scenarios of usage are envisaged by the *WikiProteins* team. In the first, experts in a given field will annotate the results of aggregated database entries and text mining. In another scenario, the user explores linked knowlets to gain understanding of a topic. Finally, *WikiProteins* can be used in "discovery mode" to find novel hypotheses revealed by exploring the network of knowlets.

The success of sister project Wikipedia, which shares links to *WikiProteins*, bodes well for a similar revolution in community knowledge development using *WikiProteins*, but it depends critically upon whether the scientific community will make the effort to participate. To succeed, *WikiProteins* needs to convince users that time spent contributing is not a donation of precious time to colleagues and competitors, but an enhancement of their own knowledge base.



Future workspaces planned for *WikiProteins*, such as WikiPeople (an intellectual networking environment), and WikiChemicals for other communities, are well underway.

WikiProteins is currently in beta trials and is accessible through <http://www.wikiprofessional.org>

Science officials

Job Hopping

In July 2009 **Andreu Mas-Colell** will assume the title of Secretary General of the European Research Council (ERC). The Spanish economist will succeed **Ernst-Ludwig Winnacker** who has held the chair since January 2007. Winnacker is packing his bags to move to Strasburg, where he will manage the Human Frontier Science Program Organization (HFSP0).

Winnacker is known as a well-versed science politician. He led the German Research Foundation (DFG) for nine years. During that time he did not shy away from speaking out on controversial issues. He consistently promoted young researchers, female scientists and interdisciplinary approaches to scientific research. We're anxiously awaiting his plans for the HFSP0.

France

CNRS Divide



France's CNRS, the largest fundamental science agency in Europe, is to be reorganised into six quasi-autonomous national institutes by the end of the year. The current CNRS-controlled departments will be placed into one of six institutes (Chemistry; Ecology and Biodiversity; Physics and Nanoscience; Nuclear and Particle Physics; Science of the Universe; Mathematics). However, life science research, which currently accounts for around 25% of the CNRS's budget, is to be treated as a new entity.

In choosing to announce the reform of the CNRS in a newspaper article (*Le Monde* 20/5/08), Valerie Pecresse, the French Minister for Higher Education and Research, shed any semblance of honouring her promises of dialogue concerning reforms "which will not be imposed from above". The scientific council of the CNRS had barely presented its own recommendations and the administrative council was to have considered the proposed reforms on 19th June. Faced with a *fait accompli*, the scientists of the CNRS can now merely note the different points, or symbolically reject them in their entirety.

Two days after Pecresse's newspaper declaration, the CNRS council met to begin discussing the plan, ahead of the formal vote on 19th June. However, a group of elected council members refused to take part in the "pretence" of discussing a decision that appears to already have been made by the government and left before the meeting could begin. "This is a profound reform of the organisation which prefigures its ultimate dissolution," a statement released by the group read. "The distinctive role of the CNRS, as a national research agency, is to explore all areas of knowledge and to coordinate interdisciplinarity. Despite a professed aim to 'decompartmentalise' research and promises to 'maintain the overall structure' of the CNRS, the ►►

►► fate of the life sciences is the prelude to a dismantlement which will prove fatal for the organisation.”

Pecresse presented more details of her plans for the life sciences on 23rd May. Currently, French life sciences are covered by four different institutions: the CNRS (fundamental biology), INSERM (medical research), INRA (agriculture) and CEA (the atomic energy commission, which does research in imagery, genomics, and certain kinds of pathology, especially relating to military applications).

Pecresse has told the four agencies to get together to find a common strategy for French life science research by the end of the year. Given the scale of their respective operations, it is in effect the CNRS and INSERM who will have to find a solution.

“Research in life sciences must be organised as a continuum which begins with fundamental work and ends in therapeutic innovation or health technology. We need the largest possible coordination to face up to international competition.”



The organisations must create means of operating together even if the government reserves the right to “intervene” in any decisions. The annual budget for public life science research in France is around €2.5 billion. Pecresse warned that they must find common ground if they are to avoid “the other solution which would consist in a fusion of the organisations”.

Rumours of such a division of the CNRS have been circulating for many months and the proposed simplified coordination of life science research is thought to be a prelude to the establishment of a new unified national life science institution under the control of the medical researchers at INSERM.

The science and higher-education trade unions and ‘Sauvons la Recherche’ (SOS Re-

search), the research movement formed in 2004 to oppose cuts in public research, have denounced the reforms as asset-stripping and a bid by the government to seize research control from the CNRS. Under the proposed changes, the CNRS will no longer appoint directors of the new institutes. Instead, they will be named by the government after an international search.

SOS Research’s pre-arranged protest march at the end of May gained extra support following reaction to Pecresse’s announcement. Inspired by the image of another persecuted minority group, SOS Research claimed that 10,000 researchers took to the streets on 27th May to proclaim their “academic pride”.

The movement threatened a total blockade of the 19th June meeting of the CNRS administrative council to prevent it from simply nodding through the reforms. It further promised to launch a series of administrative strikes similar to those which resulted in a cancellation of budget cuts by the Chirac government in 2004. -JGA-

Simpler, Faster, Cheaper, More

A Swedish team once again pushes the limits of large scale DNA sequencing.

There are many examples in the history of science demonstrating how the development of new technologies has provided the key to new insights in bioscience. The reason is simple: research problems that used to be impossible to tackle using the methods of the day, were suddenly open to investigation and analysis.

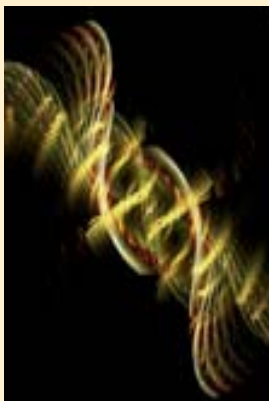
However, technology needn’t always be new to catalyse this sort of effect. Sometimes, the breakthrough is caused by making complex and expensive technology simpler, faster and cheaper to enable many labs to apply them.

Take, for example, DNA sequencing. Do you remember what, in the way of money, machines and manpower, had to be recruited to sequence the first genomes? The situation dramatically improved when, in 1998, Pål Nyrén and his colleagues from the Royal Institute of Technology in Stockholm published a new DNA sequencing method called pyrosequencing (*Science* 281: 363). Since 2003, this concept has been used to advance it into array-based pyrosequencing and has emerged as a rapid platform for relatively cheap large-scale DNA sequencing.

The latest pyrosequencing platforms generate more than 100 million nucleotide data in a seven-hour run with a single machine and each run costs about €3,500.

Another Swedish team led by Sten Linnarsson at the Department of Medical Biochemistry and Biophysics at Karolinska Institutet, may now have taken the next step forward. Their new DNA sequencing

method was recently published in *Nature Biotechnology* (vol. 26: 676-84), based on subjecting shotgun fragments to sequential hybridisation with short fluorescent probes.



The scientists split *E. coli* DNA into shotgun fragments of approximately 200 nucleotides and fixed them onto a microscope slide so that several million fragments could be analysed simultaneously. The fragments were finally rinsed in a fluid containing short DNA sequences of five nucleotides, marked with a fluorescent dye, which allowed the scientists to examine which of the short DNA sequences adhered to each fragment.

After rinsing all possible short DNA sequences over several million fragments, the scientists were then able to digitally piece the sequences together into one complete chain of the entire bacteria genome, a total of 4.5 million nucleotides.

“Everything takes place in our own specially built instrument, which comprises a microscope powerful enough to take pictures of DNA fragments, an automated pipette and a small flow chamber with a glass surface on which the reaction itself occurs,” explains Sten Linnarsson. “Moreover, the sequencing chemistry is simple, enzyme-free and consumes only dilute solutions of the probes.”

Thus, the method substantially increases speed of large scale sequencing whilst reducing the cost by about a factor of ten - perhaps the most important message for many labs.

-RN-E-

(More research results from European labs on pp. 28-33)