ECOS
SCIENCE FOR SUSTAINABILITY
www.ecosmagazine.com

Published: 22 April 2013

# Towards an open platform for digitising life on Earth

Simon Harold

**A community of scientists came together online to create a vision for the future of biodiversity research, which puts technological innovation at its heart and calls for greater openness in data sharing, standardisation and citizen science.**

Credit: Zmeel Photography/istockphoto

Your smartphone might just help us understand how the natural world works. Snap a picture of a bird, tag the image with details of where and when you took it, and you could be helping scientists to understand – and quite possibly save – the world's biodiversity.

Understanding what gives rise to the massive diversity of life on earth is perhaps the great challenge of ecological research. But to be able to predict how changes to this global system will affect the plants and creatures that live within it requires linking together huge amounts of information collected at very different scales of time and space – from microbes growing in seconds to plants competing over millennia; from nutrients cycling over nanometres to whales migrating across oceans.

Mobile technology is just one of the ways put forward to help link information and ideas about biodiversity, in an effort to ultimately understand what drives and maintains it.

An **article** recently published in *BMC Ecology*, ambitious in scope and rich with ideas, outlines a grand vision for the future of biodiversity research that seeks to answer these questions, putting openness at its heart.

Alex Hardisty from Cardiff University and Dave Roberts from London's Natural History Museum release their vision of the next decade in biodiversity research and what needs to be done to move our understanding of the natural world into the digital age.

Their white paper, written in collaboration with the Biodiversity Informatics community, was compiled through a huge community consultation effort after plans for the paper were released onto a public Google Doc. Individual contributions from 77 leaders in the biodiversity informatics field were compiled and edited by the authors to create a comprehensive vision for the future, representing a call-to-arms for funding priorities in this area.

## On complexity and connectedness

Central to this vision, the authors argue, is a need to standardise, store openly, and link together disparate datasets in a fully interoperable way. The field of genomics serves as a useful analogy in this context, with a culture of openness in data sharing, standardisation and tool development all combining to bring advances in knowledge as the cost of sequencing technologies plummet.

In contrast, the numerous tools currently available in biodiversity informatics often employ differing architectures working on differing agendas. By creating an over-arching vision, it is hoped that commonality of purpose may bring coordination toward a common goal.

A significant barrier to progress in biodiversity research is access to data. Across disciplines, only around 6-8 per cent of researchers **de deposit their data** in archives. In combination with the difficulty in recovering data from copyrights held by commercial publishers, and the widespread misunderstanding of differences between licensing and copyright, many researchers have been reluctant to release their data publicly. This issue is neatly summarised by the authors: ‘Despite considerable progress, biodiversity science is still reliant on data that is not as fully available, linkable, discoverable and accessible as it should be’.

Credit: CSIRO

Scientific institutions such as the UK’s Royal Society – no strangers to publishing volumes of biodiversity, such as the infamous **History History of Fishes** – also now call for greater openness in the digital age, and a move toward science as **an open enterprise**: ‘...although scientists do routinely exploit the massive data volumes and computing capacity of the digital age, the approach is often redolent of the paper age rather than the digital age’.

## What’s in a name?

Although some might say that taxonomy has come late to the digital party – see for example, the very recently adoption by **ICZN** and **ICI ICBM** of electronic records for species names – the opportunities now available to move species descriptions beyond the simple binary taxonomy of Linnaeus mean that taxon names may now act as access keys, unlocking information on habitats, occurrence, feeding traits, and a wealth of other information on an organism’s ecology.

The solution outlined in this vision is one where each name is assigned a persistent and unique identifier – in much the same way that genotype information is stored in repositories such as GenBank – to link information on individual species at the level of genes, all the way to whole ecosystems.

One criticism levelled at such a system is that traditional Linnaean classification may be lost, replaced by cold ID numbers. No more *Hyl Hyloscirtus princecharlesi*, no more *Agra katewinsletae* (or ***Scaptia (Plinthina) beyonceae*** – Ed). While this suggestion may risk offending practising taxonomists, the authors argue that at the rate of current species identification through these methods, the goal of achieving a comprehensive inventory of life on earth is unrealistic.

Credit: Bryan Lessard, CSIRO

Advances in molecular barcoding and metagenomics are put forward as an alternative, high-throughput solution to this problem – a technological advance to tackle a logistical conundrum. The caveat to all of this, however, is how to move computational power to the data?

Providing sufficient computational infrastructure with which to analyse the vast volumes of data that might be produced by this approach remains to be seen.

### Social media and the citizen scientist

Underlying much of this grand challenge is the potential to utilise the vast opportunities presented by modern technologies. This grand vision sees biodiversity researchers of the future becoming much better connected, with social media platforms providing an ideal venue for data and ideas to be shared efficiently.

But it is not just professional scientists who will drive these changes. Citizen scientists will be able to play a key role in enabling the vast amount of data needed to provide a system-wide perspective; encouraging future resources to be driven toward new creating new mobile technologies. Smartphone apps will be a key feature enabling amateur naturalists and enthusiastic volunteers to become tomorrow's data-gatherers.

From bird-spotters tagging species occurrences with georeferences, to armchair curators cataloguing, digitising and annotating the bewildering numbers of specimens locked away in museum vaults and private collections—the future of species identification looks likely to be greatly enhanced by the help of enthusiastic non-professionals, as access to better technology makes this ever more feasible. Crowd-sourced projects such as the **Atlas of Living Australia** will no doubt help make the massive task of cataloguing life on earth seem a little less daunting.

Credit: Steve Hillebrand /Public Domain Images **http://www.public-domain-image.com**

But is there a danger that the science might be driven by technology and not vice versa? The authors argue no, saying that the availability of new technologies allows new light to be shined on areas of research that were previously in the dark: ‘when the light changes, so does science.’

Only by making tools available to facilitate data sharing, and fostering a culture of open science in the research community, will this vision be realised in the next decade – and beyond.

In the spirit of openness, the full text from the *BMC Ecology* white paper is available under the most unrestrictive of licences, **CC0**.

*Simon Harold is Executive Editor for the BioMed Central BMC-series journals. This article is republished under Creative Commons attribution from the* **BioMed Central BMC Series blog**.

From **ECOS** online **http://www.ecosmagazine.com/?paper=EC13097**