

# **iSpot Bayesian keys: the difference that makes a difference in biological identification**

*Jon Rosewell,*

*Dept Communication & Systems, The Open University*

Identification of biological species is traditionally done by working through a paper-based dichotomous key:

Q1 Does it have six legs?

Yes go to question 2

No go to question 3.

Such keys assume either/or characters about which there can be no error and tend to produce keys that are cumbersome and error-prone in use. Dichotomous keys do not match the way people actually approach the task of identification: in practice people are more likely to note distinctive characters, to relate an unknown species to a known one, or to make a tentative identification and then check it.

We have developed computer-aided keys, based on simple Bayesian statistics, that can support these methods of working. These keys are multi-access (characters can be used in any order or skipped), use Shannon information to suggest 'best' identification characters, can give higher prior probability to commonly-encountered species or to species known to occur locally, and are robust to occasional errors. The interface supports alternative user strategies for identification such confirming or refuting tentative identification, comparing similar species, finding distinctive characters and simply browsing species descriptions. Alternative interfaces are provided for traditional computer and small-screen mobile devices.

We will outline how the Bayesian keys fit into the current iSpot.org.uk community and plans for engaging the community in developing new keys.