



## Identifiers for the Life Sciences

Estimates suggest that the Life Sciences doubles online resources every month. Species lists, character and specimen data, collectors, locations and a host of other resources abound, and so also does ambiguity.

There has been no standard way of naming, linking to, or integrating life science data. The full value of such data has not therefore been realised. Neither have we achieved a network that is far greater than the sum of its parts.

Life Science Identifiers or **LSIDs** are a new Internet standard that could address these issues.

**But don't we have identifiers already?** Yes, but identifiers like museum specimen codes and GenBank accession numbers are not by themselves very useful. You need to know that "FMNH 167095" is a specimen in the Field Museum of Natural History of Chicago. Even then, you still need to find out how to access the digital record associated with the specimen, let alone integrate diverse, yet related data.

**What about URLs?** URLs are not identifiers. URLs specify an **address**. URLs tell your web browser where a resource is, but if the resource moves, the URL breaks (the browser "404" problem). LSIDs are neatly bound to metadata, but URL's are not.

**Is there a solution?** Yes, LSIDs can be **resolved** by computers into both an address and an identifier of life science resources. LSIDs also return standardized metadata, making it easier to link disparate sources. A specimen database, for example, could use LSIDs to link specimens to scientific names in another database. A sequence database could use LSIDs to link to the specimen database.

**Can I see an example of the use of LSIDs?** Enter "*Melissotarsus insularis*" into a NCBI search and no records are found. Enter the same into the AntWeb search and you'll get a LSID "...DQ176312". Now enter this LSID into NCBI and you will get "*Melissotarsus* sp. BLF m1"; the same specimen. LSID-related applications are generating linked graphs to display and resolve many such ambiguities between different Internet resources.

Life sciences data providers can raise their profile and add value to their data by assigning and resolving LSIDs. If you want to learn more about LSIDs, see <http://jbi.nhm.ku.edu/index.php/jbi>.

Lee Belbin  
[lee@tdwg.org](mailto:lee@tdwg.org)

Rod Page  
[r.page@bio.gla.ac.uk](mailto:r.page@bio.gla.ac.uk)

Benjamin H Szekely  
[bhszekel@us.ibm.com](mailto:bhszekel@us.ibm.com)