Symbiota

May 11, 2017

Overview

Symbiota is an open source *content management system* for curating specimen- and observation-based biodiversity data. Symbiota works in conjunction with an underlying SQL database installed on a web server. Symbiota is currently utilized by hundreds of natural history collections in North America, which in turn are typically organized into 30 *thematic portals (see list below).* Jointly these portals have mobilized 20 million voucher occurrence records over the past 10 years. Symbiota portals encompass vascular plants, lichens, bryophytes, algae, fungi, invertebrates, and to a lesser degree vertebrates. Vouchered records can be linked to images, tissues, DNA sequence data, and other taxonomic information.

Why Symbiota Works

A core goal of Symbiota is to offer a platform for documenting biodiversity occurrences based on specimens and observations, by providing modular tools to share, visualize and synthesize biodiversity data for question-driven research (e.g. via checklists and interactive keys). Creating robust, multi-functional biodiversity software and data processing workflows, as well as curating data relationships for flexible management and on-line publication, are each goals capable of overwhelming the IT resources of an individual collection. A central premise of Symbiota is that through a collaborative partnership of biodiversity informaticians and collection research communities, higher quality and more publicly useful biodiversity portals can be built, rather than the old model of often highly customized stand-alone databases at an individual institution. Symbiota's open source software development framework promotes use-driven code innovation and tool creation. All of this allows collection scientists to concentrate their efforts on the curation of high-quality datasets.

A principal organizational feature of Symbiota portals is that each institutional collection is recognized separately, and controls its own data management, but all collections contribute to a general thematic framework, and are hosted together on the same server. The modular structure of portals, and ease of visualizing and exchanging data sets by end users, have made Symbiota very successful in creating large data communities and comprehensive virtual floras and faunas. The threshold for creating and maintaining a virtual collection in an established Symbiota portal is low compared to other management systems. The basic functions are reliable and easy to learn and use. Over time, their application and

Key Symbiota Design and Use Features	
Data Provider Features	**End User Features**
Open source code for relational database (SQL)	Download entire biodiversity datasets in 2-3 clicks
Portal set-up requires only 1 node with IT infrastructure and database expertise	Map georeferenced specimen records in 2-3 clicks
Shallow learning curve for most basic functions	High-resolution images and species profile pages
Allows multiple users from remote locations	Create collaborative regional taxon checklists
Built-in OCR/NLP and georeferencing tools	Interactive, region-sensitive identification keys
Powerful data/products management tools	"Name that species" educational game
Facilitates data mobilization to iDigBio, GBIF	Flash Card Quiz educational game

integration with more advanced functions can lead to powerful data products, such as DwC-A data delivery for publication and data sharing.

While Symbiota portals primarily function as web portals for collections to submit and annotate specimen or observation records, they also provide specific data packaging functions (e.g., checklist management), visualization tools (occurrence mapping and specimen images), education materials (identification keys and quizzes), and options to configure and download datasets in .csv format with 2-3 mouse clicks. Thus, Symbiota portals allow researchers to compile and screen data sets for further analysis, or educators to integrate biodiversity data for use in the classroom. By offering both live data management and periodical data snapshot support, Symbiota portals allow diverse collections to share and integrate their holdings in spite of different preferred in-house management systems. The Symbiota Working Group (SWG) provides an active user help forum and strategic community support.

Symbiota Development, Portal Services, and End-User Community

Symbiota activities can be broadly divided into **(1)** new function/module *development;* **(2)** *support services* to create, host, and expand Symbiota portal instances; and **(3)** peer-to-peer *user collaborations to promote research and education*. Because of Symbiota's open source commitment, there is no single institution or business controlling Symbiota development, so any services are either provided autonomously, or volunteered collaboratively by external parties, or offered as fee-based services under various specific arrangements. These portal support services are fundamentally independent of the Symbiota software development, but frequently a biodiversity data mobilization project may involve both and provide resources for both. The end-user community generates research and education products, provides the stimulus for new ideas and implementation of increased functionality, and develops webinars and online support documentation, including data management plans.

In 2015, the **Symbiota Working Group (SWG)** was created specifically to provide better support and functionality to existing portal communities and facilitate information and contacts to other data providers wishing to establish collections with existing portals, or create new portals. The SWG provides a linkage for outside developers interested in Symbiota, or having their applications interact with the Symbiota code and/or specific portals. The SWG coordinates and promotes the growth and sustainability of the diversifying Symbiota data portal community through increased collaborations among portals, as well as coordinating and connecting user needs, development opportunities, and services to portals. Focal SWG areas include: **(1)** sharing both expertise and training across portals; **(2)** identifying community priorities for new development and planning for implementation of added-value functions and modules; **(3)** leveraging new collaborations to seek funding for new software development and for portal growth; **(4)** providing communication and planning support to the Symbiota informatics/development community; and **(5)** improving the capacity of portal communities to provide high-quality data for both research and educational use.

Expanding Symbiota within and beyond North America

Symbiota portals have permitted many 100s of North American biodiversity collections to share their holdings on-line. The SWG is ready to facilitate information, contacts, and where feasible, direct support to any collection in the world interested in exploring and promoting Symbiota and creating portal instances. Our primary motivation for reaching out to institutions and collections is that using a common Symbiota portal is a great way to promote regional collaborations in research and education among collections within a portal, and ultimately global-scale collaboration among portals. Portals can be organized by taxonomic interest (e.g., vascular plants) or by institute that include all taxa. If you are interested in knowing more about Symbiota, please contact SWG lead coordinators Neil Cobb (Neil.Cobb@nau.edu) or Andrew Miller (amiller7@illinois.edu).

Key Symbiota Websites

Homepage: http://symbiota.org/
Code @ GitHub: https://github.com/Symbiota
Citable publication: http://bdj.pensoft.net/articles.php?id=1114
Google Group (support): http://symbiota.org/docs/google-group/
Symbiota Working Group: https://www.idigbio.org/wiki/index.php/Symbiota_Working_Group
Symbiota Calendar: Google Calendar

List of Known Symbiota Portals (Updated February, 2016)

Portal Name & Theme	Portal Web Address (URL)
Consortium of North American Lichen Herbaria	http://lichenportal.org/portal/index.php
Arctic Lichen Flora	http://lichenportal.org/arctic/index.php
Consortium of North American Bryophyte Herbaria	http://bryophyteportal.org/portal/
Frullania Collaborative Research Network	http://bryophyteportal.org/frullania/
Macroalgal Consortium Herbarium Portal	http://macroalgae.org/portal/index.php
MyCoPortal	http://mycoportal.org/portal/index.php
Smithsonian Tropical Research Institute Portal (STRI)	http://stricollections.org/portal/
Aquatic Invasives	http://greatlakesinvasives.org/portal/index.php
Consortium of Midwest Herbaria	http://midwestherbaria.org/portal/index.php
SEINet (Original Portal – Plants)	http://swbiodiversity.org/seinet/index.php
Intermountain Region Herbarium Network (IRHN)	http://intermountainbiota.org/portal/index.php
SouthEast Regional Network of Expertise and Collections (SERNEC)	http://sernecportal.org/portal/
North American Network of Small Herbaria	http://nansh.org/portal/index.php
Northern Great Plains Herbaria	http://ngpherbaria.org/portal/index.php
New Mexico Biodiversity Portal	http://swbiodiversity.unm.edu/
Consortium of Northeastern Herbaria (CNH)	http://portal.neherbaria.org/portal/
Madrean Archipelago Biodiversity Assessment (MABA) Flora	http://madrean.org/symbflora/projects/index.php
Madrean Archipelago Biodiversity Assessment (MABA) Fauna	http://madrean.org/symbfauna/projects/index.php
Herbario Virtual Austral Americano	http://herbariovaa.org/
CoTRAM – Cooperative Taxonomic Resource for Amer. Myrtaceae	http://cotram.org/
InvertEBase Data Portal	http://invertebase.org/portal/index.php
Symbiota Collections of Arthropods Network (SCAN)	http://symbiota4.acis.ufl.edu/scan/portal/index.php
Neotropical Entomology	http://symbiota.org/neotrop/entomology/index.php
Neotropical Flora	http://symbiota.org/neotrop/plantae/index.php
Monarch (California Academy of Sciences)	http://monarch.calacademy.org/
The Lundell Plant Diversity Portal	http://prc-symbiota.tacc.utexas.edu/index.php
Virtual Flora of Wisconsin	http://symbiota.botany.wisc.edu
Red de Herbarios del Noroeste de México	http://herbanwmex.net/portal/index.php
University of Colorado Herbarium	https://botanydb.colorado.edu/
Open Herbarium	http://openherbarium.org/index.php
Consortium of Pacific Herbaria	http://www.pacificherbaria.org/