

El Atlas de la Biodiversidad de Costa Rica (CRBio)

María Auxiliadora Mora, Manuel Vargas

Instituto Nacional de Biodiversidad de Costa Rica (INBio)
mmora@inbio.ac.cr, vargas@inbio.ac.cr

Resumen: El Atlas de la Biodiversidad de Costa Rica (CRBio) fue creado en 2006 con la misión de proveer acceso integrado, libre y gratuito a los datos y a la información de la biodiversidad costarricense para apoyar la ciencia, la educación y la conservación de la biodiversidad. El portal en Internet de CRBio (<http://www.crbio.cr/>) fue renovado en 2016 con componentes de software libre desarrollados por el Atlas de la Biodiversidad de Australia (ALA), los cuales han sido utilizados también en otros países que han conformado una comunidad de programadores y usuarios con el apoyo de la Infraestructura Mundial de Información en Biodiversidad (GBIF) y de ALA. El nuevo portal de CRBio incluye 5.000 páginas de especies, más de 50.000 imágenes y casi siete millones de registros de presencia de especies georreferenciados en Costa Rica. Estos registros provienen de más de 800 bases de datos publicadas por centros de investigación de 33 países. Los diferentes tipos de datos e información contenidos en CRBio siguen los estándares definidos por Estándares de Información en Biodiversidad (TDWG), tales como el Darwin Core Archive para los registros de presencia de especies y el Plinian Core para las páginas de especies. El empleo de estos estándares y de herramientas como la Herramienta de Publicación de Datos (IPT), por parte de los publicadores, facilita compartir e intercambiar los datos de CRBio con iniciativas globales de informática de la biodiversidad como GBIF, la Enciclopedia de la Vida (EOL) y la Biblioteca del Patrimonio de la Biodiversidad (BHL).

Abstract: The Atlas of Living Costa Rica (CRBio) was launched in 2006 aiming at providing integrated, free, and open access to data and information about the Costa Rican biodiversity to support science, education, and biodiversity conservation. The CRBio data portal (<http://www.crbio.cr/>) was updated in 2016 with free software components developed by the Atlas of Living Australia (ALA), which have been also used in other countries who have established a community of developers and users supported by the Global Biodiversity Information Facility (GBIF) and ALA. The new data portal of CRBio includes 5,000 species pages, more than 50,000 images, and around seven million georeferenced occurrence records. The occurrence records are part of more than 800 databases and were published by research centers located in 33 countries. The data and information aggregated by CRBio follow standards developed by the Biodiversity Information Standards (TDWG), such as Darwin Core Archive for occurrence records and Plinian Core for species pages. The use of these standards and publishing tools, like the Integrated Publishing Toolkit (IPT), facilitates sharing and exchanging data with global Biodiversity Informatics initiatives such as GBIF, the Encyclopedia of Life (EOL) and the Biodiversity Heritage Library (BHL).

Introducción

El Atlas de la Biodiversidad de Costa Rica (CRBio) fue creado en 2006 con la misión de proveer acceso integrado, libre y gratuito a los datos y a la información de la biodiversidad costarricense para apoyar la ciencia, la educación y la conservación de la

biodiversidad. El portal en Internet de CRBio agrupa datos e información generados por instituciones nacionales e internacionales que se dedican a la investigación, la educación, el uso sostenible y la gestión de la biodiversidad. Páginas de especies, registros de presencia (ej. especímenes, observaciones), capas geográficas (ej. división política, áreas silvestres protegidas), nomenclatura científica (ej. nombres, clasificaciones), información bibliográfica y datos multimedia, están disponibles en formatos estandarizados como el Darwin Core (<http://rs.tdwg.org/dwc/>) y el Plinian Core (<https://github.com/PlinianCore>), ambos definidos por Estándares de Información en Biodiversidad (TDWG), la cual es una organización científica, educacional y sin fines de lucro que trabaja en el desarrollo de los estándares abiertos para el intercambio de información sobre biodiversidad empleados por las principales iniciativas a nivel mundial.

CRBio utiliza también software libre desarrollado por la comunidad internacional de informática de la biodiversidad. Por ejemplo, inicialmente usó el protocolo y las herramientas DiGIR (*Distributed Generic Information Retrieval*, <http://digir.sourceforge.net/>) para implementar el portal en Internet y los nodos de datos de la red; en una etapa posterior el portal se actualizó con software desarrollado por la Infraestructura Mundial de Información en Biodiversidad (GBIF) y los nodos fueron implementados utilizando el protocolo TAPIR (*TDWG Access Protocol for Information Retrieval*, <http://www.tdwg.org/activities/tapir/>). En 2016, gracias a un proyecto ejecutado por el Instituto Nacional de Biodiversidad (INBio) y patrocinado por el Consejo Nacional de Investigaciones Científicas y Tecnológicas (Conicit) y el Ministerio de Ciencia, Tecnología y Telecomunicaciones (Micitt), el portal en Internet de CRBio (<http://www.crbio.cr/>) fue renovado con componentes de software desarrollados por el Atlas de la Biodiversidad de Australia (ALA), los cuales han sido instalados y adaptados también en otros países (ej. Argentina, Brasil, Escocia, España, Francia, Portugal, Reino Unido) que han conformado una comunidad de programadores y usuarios con el apoyo de GBIF y de ALA. Para compartir los datos, los nodos utilizan actualmente la Herramienta de Publicación de Datos (IPT), la cual es la recomendada por GBIF y otras iniciativas globales de informática de la biodiversidad. El nuevo portal de CRBio incluye 5.000 páginas de especies, más de 50.000 imágenes y casi siete millones de registros de presencia de especies georreferenciados en Costa Rica. Estos registros provienen de cerca de 800 bases de datos publicadas por centros de investigación de más de 33 países. La figura 1 muestra los sitios de recolección u observación de especies disponibles en el portal de CRBio.

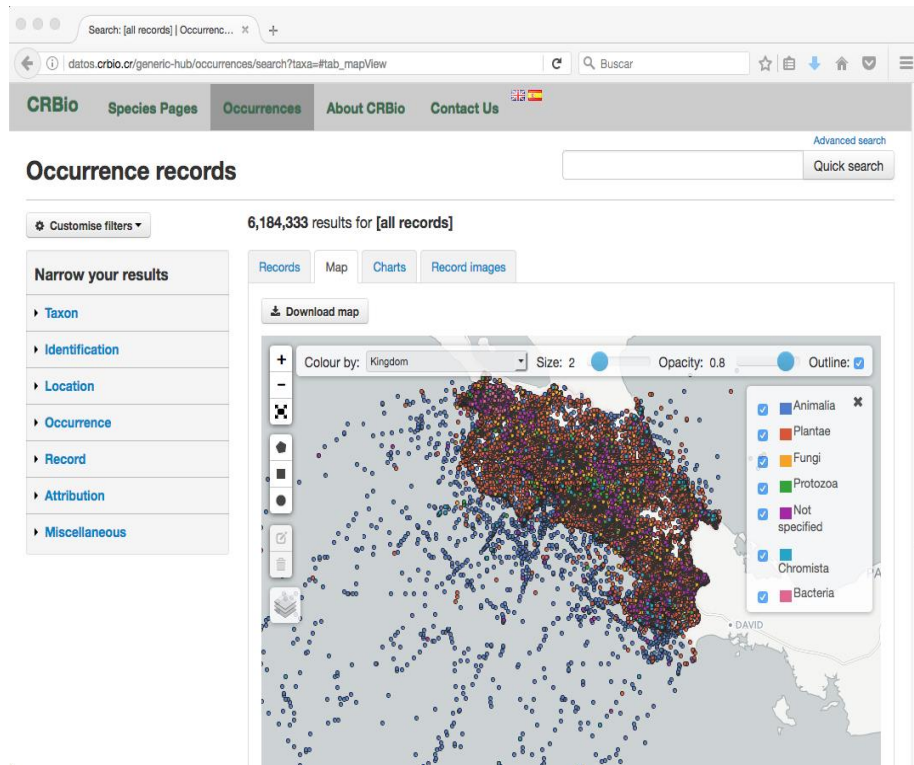


Figura 1. CRBio es un esfuerzo colaborativo que integra cerca de 7 millones de registros de presencia de especies de la biodiversidad de Costa Rica, publicados por centros de investigación de más de 30 países.

CRBio es un esfuerzo colaborativo que ha contado con la participación de instituciones como el INBio, el Área de Conservación Guanacaste (ACG), la Organización para Estudios Tropicales (OET), el Sistema Nacional de Áreas de Conservación (Sinac), la Comisión Nacional para la Gestión de la Biodiversidad (Conagebio), la Escuela de Biología de la Universidad de Costa Rica (UCR), el Museo Nacional de Costa Rica (MNCR), el Laboratorio de Recursos Naturales y Vida Silvestre (Lamavisi) de la Universidad Nacional de Costa Rica (UNA) y los Observatorios de Aves de Costa Rica (CRBO).

Impacto de la investigación desarrollada

CRBio ha realizado una contribución importante a Costa Rica y a la comunidad internacional de informática de la biodiversidad. A nivel nacional, CRBio ha llenado un vacío de integración de fuentes de datos sobre biodiversidad y publicación de información por medio de un mecanismo de acceso centralizado, libre y gratuito. Asimismo, ha propiciado la comunicación entre instituciones nacionales relacionadas con el manejo de información sobre biodiversidad. También ha fortalecido la capacidad informática en el recurso humano nacional para capturar, curar, analizar y publicar datos e información sobre la biodiversidad de Costa Rica.

A nivel internacional, se ha propiciado la formación y consolidación de una comunidad de desarrolladores relacionados con las herramientas de ALA y de aplicaciones de informática de la biodiversidad en general. Se ha contribuido también al fortalecimiento de capacidades en informática de la biodiversidad por medio de proyectos de tutoría orientados a compartir el conocimiento y la experiencia con instituciones en otros países (ej. Nicaragua, Argentina, Perú, Chile, Cuba, Bután y Benín). Estos proyectos fueron financiados por GBIF, la Fundación de Biodiversidad JRS y el Programa de Cooperación Sur-Sur de Fundecooperación con fondos del Reino de los Países Bajos.

Red internacional de colaboración

Como se ha mencionado, varios países y organizaciones han reutilizado y adaptado los componentes de software desarrollados por ALA para sus portales de datos de biodiversidad, lo que ha dado lugar a la formación de una comunidad de usuarios y programadores interesados en el mantenimiento y la mejora de estas herramientas.

Esta comunidad comenzó a formarse en Costa Rica en el año 2013, cuando GBIF patrocinó un proyecto de fortalecimiento al desarrollo de capacidades entre el nodo costarricense y el nodo australiano, el cual estuvo enfocado a estudiar el software del portal de ALA. Se organizaron dos talleres en Costa Rica a los cuales asistieron especialistas de Australia, Costa Rica, España, México y Brasil. Varios talleres adicionales se han llevado a cabo desde entonces en Australia, España, Francia y Argentina, la mayoría financiados por GBIF y también con el apoyo de otros organismos, como la Unión Europea. Además de estos eventos presenciales, los cuales se procuran organizar como mínimo una vez al año, la comunidad se comunica a través de listas de correo, grupos de *chat* y facilidades que ofrece GitHub (<https://github.com/>) para el desarrollo colaborativo de software.

ALA comparte el código fuente y la documentación de su software en el repositorio GitHub (<https://github.com/AtlasOfLivingAustralia/>) con licencias de software libre, lo que permite que los interesados extiendan y compartan a su vez estos recursos. Entre los proyectos derivados de este esquema de colaboración pueden mencionarse los portales de datos de los nodos GBIF de Argentina (<http://datos.sndb.mincyt.gob.ar/>), España (<http://datos.gbif.es/>), Francia (<http://portail.gbif.fr/>) y Portugal (<http://dados.gbif.pt/>), así como el Atlas de la Red Nacional de Biodiversidad de

Escocia (<https://scotland.nbnatlas.org/>) y el portal de datos del Instituto Chico Mendes, en Brasil (<https://portaldabiodiversidade.icmbio.gov.br/portal/>). Los nodos GBIF de Alemania, Suecia, México y Canadensys (organización adscrita a la Universidad de Montreal), también están implementando portales de datos con base en estas tecnologías, lo que demuestra el creciente interés en las aplicaciones desarrolladas por ALA y sus colaboradores, así como el potencial que tienen los datos abiertos y el software libre en la emergente área de conocimiento de la informática de la biodiversidad.

Tecnologías de información y comunicación empleadas en el desarrollo de la investigación

Los diferentes tipos de datos e información contenidos en CRBio siguen los estándares definidos por TDWG, tales como el *Darwin Core Archive* para los registros de presencia de especies y el *Plinian Core* para los resúmenes de información general sobre especies.

El portal en Internet está basado en software libre como:

- El Sistema de Información Atta: software desarrollado por INBio para administrar la información sobre la biodiversidad de Costa Rica que la institución genera. El sistema incluye un módulo para la publicación de páginas de especies que también está en uso en CRBio.
- El gestor de contenidos WordPress (<https://wordpress.com/>).
- Módulos del portal de ALA para el manejo de:
 - datos de presencia de especies y
 - metadatos de instituciones y colecciones.

La figura 2 muestra la arquitectura del portal de ALA. Los recuadros en rojo representan la funcionalidad utilizada por CRBio. La arquitectura está dividida en capas: en la base de la figura se encuentra la capa de preprocesamiento de datos (*offline processing*), compuesta por procesos efectuados antes de incorporar los datos al portal para asegurar el cumplimiento de los estándares y mejorar la calidad de estos; la capa inmediatamente superior (*DBs, Indexes, Filesystem storage*) incluye los mecanismos de persistencia utilizados por el sistema; la siguiente capa es la de servicios web (*Web services*), que permite que cierta funcionalidad del portal sea utilizada por otras aplicaciones y finalmente en la capa superior, se encuentra la funcionalidad de consulta y visualización de los datos (*Front End Apps*).

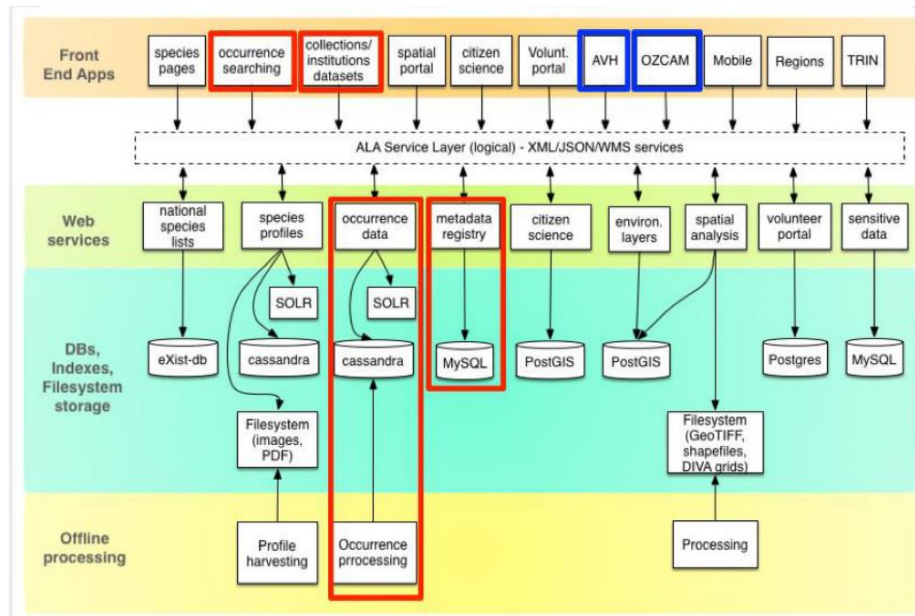


Figura 2. Arquitectura del portal desarrollado por el Atlas de Biodiversidad de Australia (ALA) en uso en CRBio desde el 2016.