

Darwin Core: Estándar para la Gestión de Datos Biológicos Primarios en la UTN

Sania Ortega Andrade^a, Alexander Guevara Vega^b

^a Universidad Técnica del Norte, Docente de la Facultad de Ingeniería en Ciencias
Agropecuarias y Ambientales, FICAYA, Av. 17 de Julio, Ibarra, Ecuador
smortega@utn.edu.ec

^b Universidad Técnica del Norte, Docente de la Facultad de Ingeniería en Ciencias Aplicadas,
FICA, Av. 17 de Julio, Ibarra, Ecuador
alexguevara@utn.edu.ec

Resumen. El incremento de investigaciones biológicas y de biodiversidad en los centros universitarios constituye una fuente importante de datos que generan conocimiento, sin embargo, no cuentan con estándares ni una plataforma para la gestión de datos biológicos primarios. A nivel nacional es perceptible la falta de integración y socialización con otras universidades, redes de ONG's, museos, herbarios u otras instituciones nacionales que faciliten el intercambio de datos y conocimientos. Esto ha generado que las fuentes de información del Ecuador y sus datos no sean difundidas en su totalidad. Por lo tanto, el objetivo de este proyecto es desarrollar una plataforma tecnológica basada en Darwin Core (DwC) como estándar, que optimice la gestión de datos biológicos primarios en los laboratorios de Investigación Biológica y Ambiental de la Facultad de Ingeniería en Ciencias Agropecuarias y Ambientales (FICAYA) de la Universidad Técnica del Norte (UTN). El proceso de análisis de los datos biológicos se consolida de manera sistémica, a través de procesos, actividades y herramientas que, al ser agrupadas en fases, permiten obtener un repositorio estandarizado. Para esto se utilizó la metodología de SCRUM la cual guio de forma ágil el proceso de desarrollo de la plataforma tecnológica; además se aplicó DwC como estándar para la divulgación e intercambio de información biológica. La plataforma esta estandarizado acorde a los requerimientos y procesos de cada área biológica lo que permitió el fortalecimiento de la publicación y gestión de datos, además su uso generalizado constituirá positivamente para la toma de decisiones en pro de la conservación ambiental.

Palabras Clave: Dato, colecciones, biodiversidad, Darwin Core, Web, sistemática, taxonomía.

Abstract. The increase in the number of biological and biodiversity studies in university centers constitutes an important source of data that generates knowledge, however, these centers do not yet have either standards or a platform for the management of primary biological data. At the national level, the lack of integration and communication with other universities, NGO networks, museums, herbariums or other national institutions that facilitate the exchange of data and knowledge is perceptible. This result of this lack of data exchange is that the sources of information from Ecuador and its data have not been completely shared. Therefore, the objective of this project was to develop a technological platform based on Darwin Core (DwC) as a standard, which would optimize the management of primary biological data in the laboratories of Biological and Environmental Investigation in the Faculty of Environmental and Agricultural Sciences (FICAYA) of the Universidad Técnica del Norte (UTN). The biological data analysis process is consolidated systematically by means of processes, activities and tools that, grouped in phases, build a standardized data repository. The SCRUM methodology was used to guide the agile process of development of the technological platform; additionally, DwC was applied as a standard for sharing and interchange of biological information. The platform is standardized

according to the requirements of each biological area which strengthened the publication and management of data, additionally, its generalized use will positively influence decision-making in favor of environmental conservation.

Key words: Data, collections, biodiversity, Darwin Core, Web, systematic, taxonomy.

1 Introducción

El incremento de procesos e iniciativas de la investigación referente a la conservación de la biodiversidad en las Instituciones de Educación Superior (IES) e Institutos para la investigación Biológica en el Ecuador constituye una fuente importante de generación de datos, información y conocimiento. La UTN y la FICAYA se han convertido en un referente a nivel de la región norte del Ecuador en la generación de estudios de biodiversidad, conservación de recursos renovables, estudios de prospección biológica, colección y caracterización de especies biológicas de distintos grupos taxonómicos, entre ellos animales, plantas, líquenes, hongos y microorganismos del Ecuador y la Antártida, todo este catálogo de contenidos son ingresados en documentos informativos, tesis, bases de datos individuales en Excel, entre otros tipos de archivos, mismos que tienen la necesidad de ser publicados, difundidos y sobre todo reutilizados como recursos de acceso abierto.

El avance de la tecnología ha permitido pasar de libros de campo a bases de datos como archivos planos de Excel y Access dentro de centros académicos biológicos, museos o institutos. Hoy en la actualidad ese avance ha generado una transformación digital e inmersión de contenidos multimedia como base de datos estructuradas y geoespaciales, así como nuevos espacios de visualización de contenidos biológicos en 2D y 3D, las cuales son fundamentales para la construcción de un ecosistema dinámico de información biológica enriquecida.

La falta de procesos bien definidos en los laboratorios de investigación biológica y ambiental de la FICAYA como normas, estándares y protocolos no han facilitado el registro y publicación de datos biológicos de una forma adecuada. El incremento de investigaciones en estas áreas prevé un aumento de estos datos, lo cual dificultará aún más este proceso, con un alto grado de centralización dependiente de un bajo número de digitadores, investigadores y estudiantes. El problema identificado trae como consecuencia cuellos de botella en el proceso de registro de especies al utilizar un solo documento para el registro, acumulación de trabajo, pérdida de información, inseguridad, inconsistencia, duplicidad e integridad de la información y con ello limitaciones técnicas en las consultas simultáneas de la información, creando barreras de adopción, publicación e intercambio de información biológica.

El presente trabajo es sobre el desarrollo e implementación de una Plataforma Tecnológica Web basada en la aplicación de DwC como estándar para la Gestión de datos Biológicos Primarios en la UTN, para generar una red de colaboración entre docentes, estudiantes y comunidad científica, definiendo nuevos métodos y modelos para la gestión de datos biológicos-basado en estándares y protocolos para el intercambio y la divulgación.

2 Impacto

En la UTN, cada uno de los programas y proyectos de investigación biológica y ambiental se manejaban de manera independiente, de tal forma que generaban gran cantidad de datos almacenados a través de hojas de caculo y herramientas no estructuradas, las cuales se convertían en un problema al momento de registrar y publicar, por lo que no se lograba obtener los resultados y beneficios esperados, como: consolidar la integración, colaboración y socialización para establecer redes de investigación que permitan el intercambio de conocimientos; el enriquecimiento de la base de datos y el diálogo de saberes sobre el bioconocimiento.

Técnicamente el proyecto es importante porque desarrollo la plataforma tecnológica Web mediante la aplicación de DwC como estándar en los laboratorios de investigación biológica y ambiental de la FICAYA-UTN, y mejoró la gestión de datos biológicos primarios, optimizando el proceso de registro y publicación de datos; permitiendo el acceso en tiempo real a un mayor número de personas.

Económicamente el proyecto redujo costos a la universidad, en la obtención de la información, optimización de las tareas por parte de los usuarios y en alcanzar una gestión íntegramente computarizada dentro y fuera de la universidad, ya que el mantenimiento y curaduría de colecciones biológicas conlleva gastos altos en infraestructura y manejo.

Socialmente está asociado a dos pilares estratégicos de la UTN, Investigación y Vinculación y al Plan del Buen Vivir que el Gobierno del Ecuador está impulsando para brindar soluciones a las problemáticas que la sociedad tiene al no contar con información biológica y ambiental del patrimonio natural. La mayoría de centros vinculados a la biodiversidad en el Ecuador tienen accesos restringidos a los datos y la difusión es interna, creando barreras que hoy en día la tecnología está tratando de minimizar gracias a Iniciativas de Open Access.

Con la creciente ola tecnológica y la divulgación de información científica, la gestión de datos biológicos primarios permitió administrar, analizar, almacenar y publicar datos taxonómicos parametrizados con estándares y normas de publicación e intercambio de información biológica, minimizando las barreras de adopción y maximizando la reutilización.

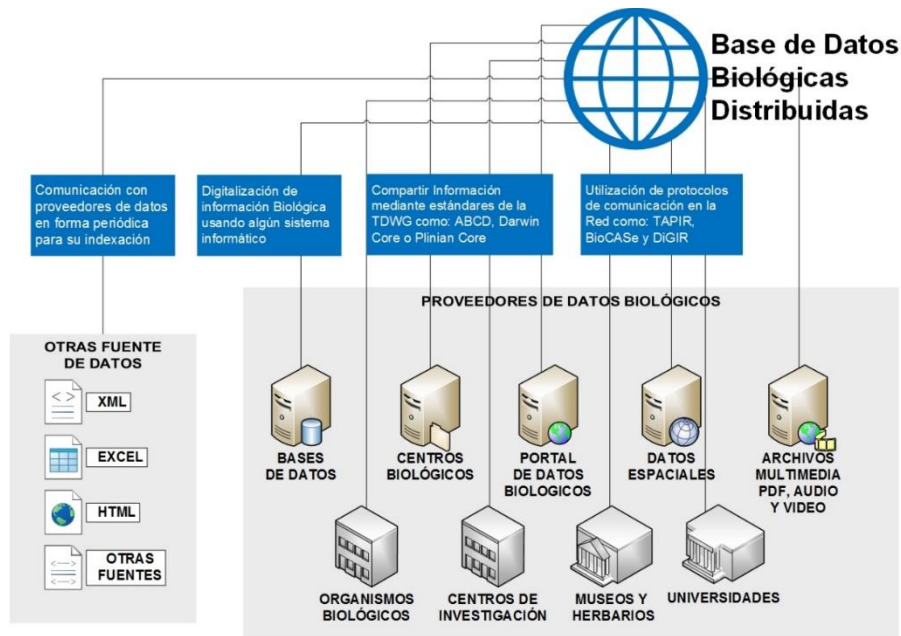


Figura 1. Esquema macro de la integración de agentes internos y externos en la plataforma tecnológica.

Bajo este contexto para la investigación e implementación de la plataforma tecnológica se tomó como línea base datos biológicos basados en su contenido construyendo un repositorio de datos estructurales primarios sometidos por la comunidad científica, esta base de datos es fundamentales para la construcción de un sistema de información biológica en la zona norte del Ecuador.

3 TIC aplicadas al proyecto

Para el desarrollo del SIGBIO, se realizó un conversatorio previo con los técnicos involucrados del Área de Gestión Web (AGWEB) de la Dirección de Desarrollo Tecnológico e Informático (DDTI) de la UTN y el equipo de investigación de la FICAYA, donde se acordó aprovechar los recursos existentes y la infraestructura actual que mantiene la UTN como base tecnológica para diseñar la siguiente arquitectura tecnológica:

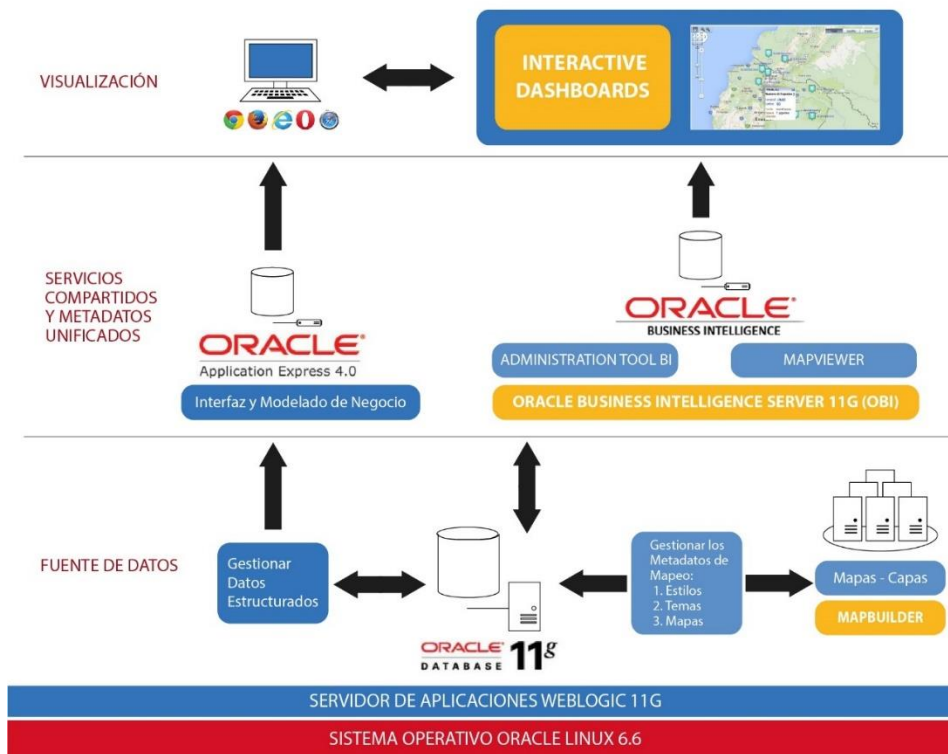
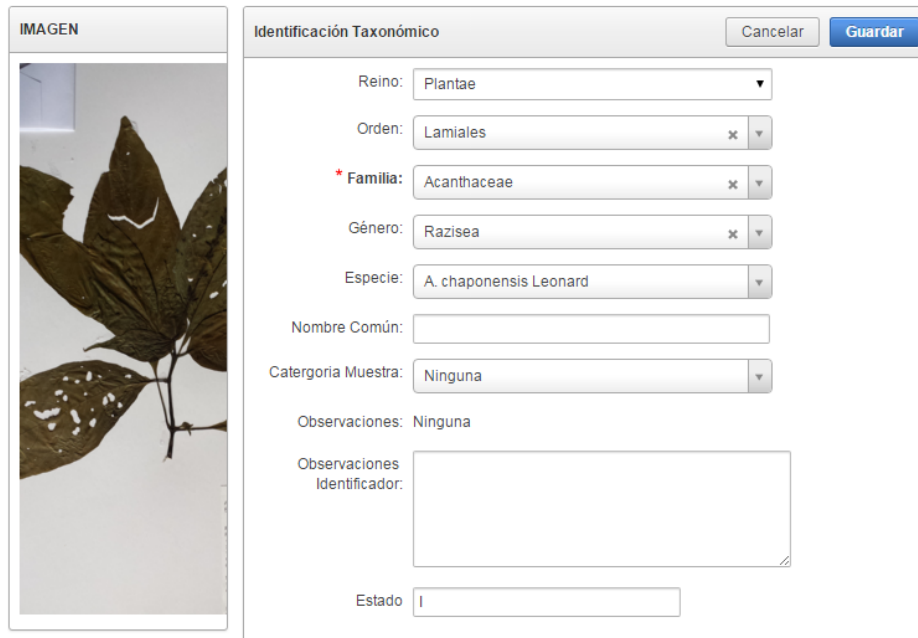


Figura 2. Arquitectura Tecnológica de la plataforma Web para la gestión de datos biológicos primarios.



The screenshot shows a web interface for biological data management. On the left, there is a section labeled 'IMAGEN' containing a photograph of a leaf with several holes, likely from a cicada. On the right, there is a form titled 'Identificación Taxonómico' with a 'Cancelar' button and a 'Guardar' button. The form contains the following fields:

- Reino:
- Orden:
- * Familia:
- Género:
- Especie:
- Nombre Común:
- Categoría Muestra:
- Observaciones:
- Observaciones Identificador:
- Estado:

Figura 3. Módulo de Gestión biológica ingreso de información taxonómica en la plataforma tecnológica Web para la gestión de datos biológicos primarios.

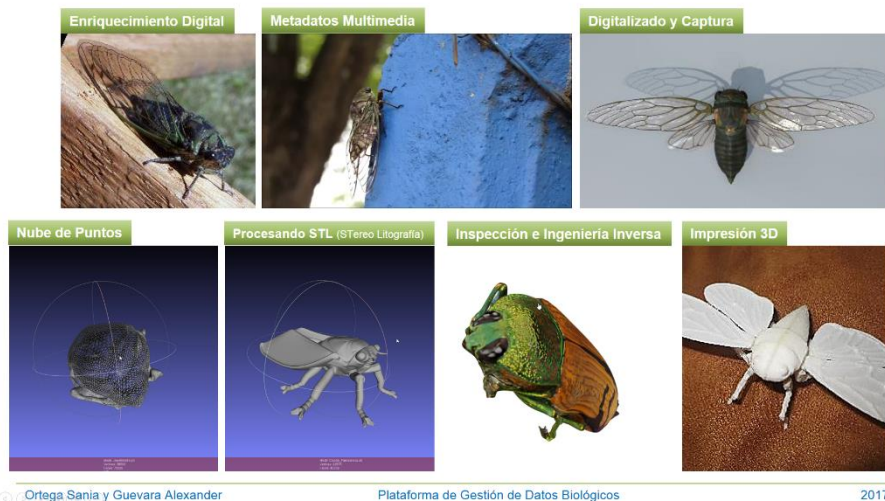


Figura 4. Enriquecimiento digital en el módulo de Gestión biológica visualizados en la plataforma tecnológica Web para la gestión de datos biológicos primarios.

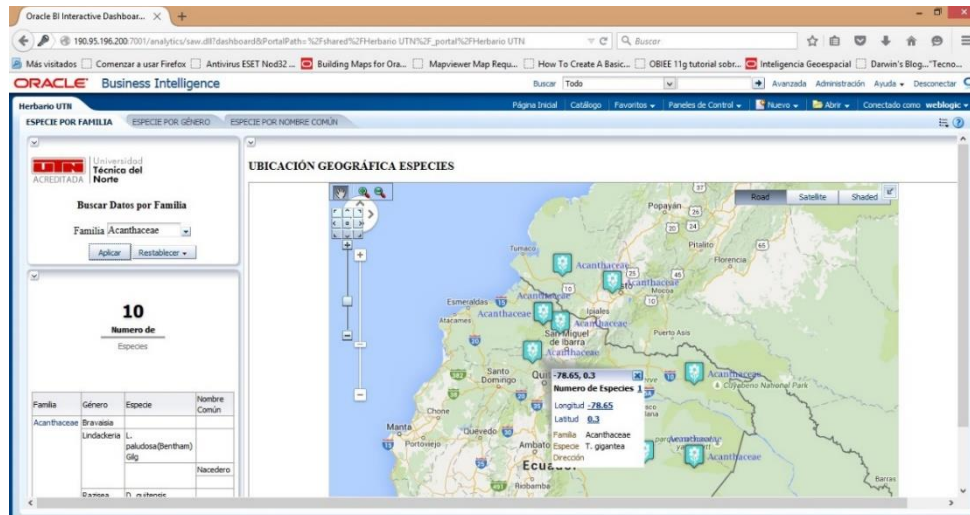


Figura 5. Módulo de GIS de la plataforma tecnológica Web para la gestión de datos biológicos primarios.

4 Resultados

Uno de los resultados principales fue la creación de un mapa de procesos para la gestión de datos biológicos en los laboratorios de investigación biológica y ambiental de la FICAYA de la UTN, definiendo una jerarquía de procesos y procedimiento orientados a los esquemas institucionales.

Al implementar la plataforma tecnológica Web se pudo mejorar la gestión de datos biológicos primarios lo que permitió el intercambio de conocimientos, el enriquecimiento de la base de datos y el diálogo de saberes sobre el bioconocimiento; aumentó notablemente la centralización de la información, facilitando el registro y publicación de los datos en tiempo real y de acceso abierto beneficiando a la sociedad en general.

Incrementó la divulgación de la información biológica para dar a conocer la importancia que tiene el conocimiento de la diversidad de la flora, fauna y microorganismos a nivel nacional y sus valores científicos y económicos para los investigadores, docentes, estudiantes y público en general.

Fomentó la implementación de procesos estandarizados en todo el ciclo de vida para la gestión de los datos biológicos.

Automatizó el proceso de etiquetación del material biológico y generación de sus respectivas etiquetas (código de barras y/o códigos QR).

Almacena la información del material biológico en la base de datos de la plataforma WEB mismo que se mapean con los términos de DwC como estándar y así compartir y ser consultadas simultáneamente con otros recursos de información sobre biodiversidad.

Obtención de información de datos biológicos sobre especímenes recolectados: catálogos, colecciones, especies y etnobiología.

5 Alianzas y redes de colaboración

El desarrollo de esta investigación permitió establecer alianzas estratégicas entre la UTN y Red de Investigación en Biodiversidad del Ecuador (RIBIODE) constituida y registrada el 14 de noviembre de 2014 con resolución CRCPEZ1-S07 N°. 003-14 a cargo del Comité Regional consultivo de Planificación de la Educación Superior Zona 1 del Ecuador, con el fin de garantizar el libre acceso a la información no solo de la universidad sino de todos los institutos de investigación biológica involucrados en la gestión y conservación de la Biodiversidad.

Un acercamiento preliminar entre la Universidad Regional Amazónica (IKIAM) y el Instituto Nacional de Biodiversidad del Ecuador (INABIO) ha permitido la divulgación de esta investigación y futuras alianzas para el intercambio y la gestión de los datos biológicos a nivel Nacional (Múltiples gestores en un solo sistema.).

6 Experiencias aprendidas

Casa adentro la multidisciplinariedad permitió el trabajo conjunto entre la Dirección de Desarrollo Tecnológico e Informático (DDTI) de la UTN y el equipo de investigación de la FICAYA, juntos fusionaron sus conocimientos para la creación de un sistema para la gestión de datos biológicos primarios, sino también para la esquematización de procesos y seguimiento del ciclo de vida de los datos.

Un aspecto importante para engrandecer un proceso es la divulgación de las investigaciones y el entendimiento entre los investigadores involucrados en el proceso, con los líderes que gestionan políticas, normas y reglamentos. El Ecuador es un país Megadiverso, pero no cuenta con un sistema eficiente para gestionar sus datos. Aunar esfuerzos y establecer alianzas para compartir el conocimiento es la única opción para lograr la gestión adecuada.