

Título: QUALITY AND FITNESS-FOR-USE ASSESSMENTS ON THE PRIMARY DATA INDEXED AT THE GLOBAL BIODIVERSITY INFORMATION FACILITY (GBIF)

Nombre: Otegui Tellechea, Javier

Universidad: NAVARRA

Departamento: ZOOLOGIA Y ECOLOGIA

Fecha de lectura: 28/06/2012

Dirección:

> **Director:** Arturo Hugo Ariño Plana

Tribunal:

> **presidente:** RAFAEL JORDANA BUTTICAZ

> **secretario:** Ricardo Ibáñez Gastón

> **vocal:** LUIS VILLAR PEREZ

> **vocal:** Vishwas Cahvan

> **vocal:** Juan Junoy Pintos

Descriptorios:

> BASES DE DATOS

> SISTEMAS AUTOMATIZADOS DE CONTROL DE CALIDAD

El fichero de tesis no ha sido incorporado al sistema.

Resumen: Aunque la biodiversidad siempre ha sufrido diversas amenazas debido a su naturaleza dinámica, en las últimas décadas se ha llegado, principalmente por el efecto del creciente impacto humano en el medio ambiente, a un escenario que se conoce como la *¿crisis de la biodiversidad¿*, una crisis cuyo tratamiento, se dice, depende en buena medida de la ampliación del conocimiento sobre biodiversidad. Como parte fundamental de la solución para esta demanda de conocimiento, se crea la Infraestructura Global de Información de Biodiversidad (GBIF por sus siglas en inglés) con el fin de *¿hacer disponibles de manera libre y universal los datos de biodiversidad a través de Internet¿*. Hoy por hoy, GBIF representa la mayor red de instituciones editoras de datos primarios de biodiversidad (PBD). Sin embargo, se ha defendido progresivamente entre los implicados en el estudio de la biodiversidad la conveniencia de evaluar la calidad de los PBDs servidos a través de GBIF de manera que los proveedores *¿dueños de la propiedad intelectual de los mismos ¿* sean conscientes del estado de los registros que comparten. En este marco se encuadra la presente Tesis Doctoral, que ha tenido como objetivo general contribuir al desarrollo de la informática aplicada a la biodiversidad. Esto se traduce en los siguientes objetivos particulares: (a) evaluar la calidad y usabilidad de los PBDs accesibles a través de la red de instituciones editoras de datos de GBIF; (b) analizar con mayor profundidad desde distintas perspectivas aspectos más detallados del conjunto de PBDs disponibles para generar recomendaciones de mejora; (c) facilitar a la comunidad el acceder a distintos análisis sencillos de calidad de datos mediante el desarrollo de una herramienta web de visualización de colecciones de PBDs; y (d) aplicar el conocimiento adquirido para realizar estudios ejemplares del uso de los datos disponibles en la red de

GBIF. Los resultados muestran que, si bien, la compleción de los registros en general es buena, la distribución de sus componentes no lo es tanto, permitiendo detectar de la extracción de sus pautas el efecto *¿low-hanging fruit¿* (la sobrerrepresentación de los registros sencillos de obtener en relación a los más complicados). El gran volumen de datos disponibles no merma la necesidad de aumentar la información disponible, bien *¿descubriendo¿* a la red colecciones existentes o continuando con el muestreo de zonas de elevada biodiversidad con el fin de aumentar la usabilidad de la información.