

# CONJUNTOS DE DATOS REPATRIADOS 2018

**eBird, iNaturalist, GBIF y Xeno-canto**

**Diciembre 2018**

*Versión - 1.0*

---

**Cítese como:** SiB Colombia (2018). *Conjuntos de datos repatriados 2018 - eBird, iNaturalist, GBIF y Xeno-canto*. Sistema de Información sobre Biodiversidad de Colombia, Bogotá D.C., Colombia, 7 pp.

---

**Licencia:** Este documento se publica bajo una licencia *Creative Commons Attribution 4.0*



**Control del documento:**

<b>Versión</b>	<b>Descripción</b>	<b>Fecha publicación</b>	<b>Autor(es)</b>
1.0	Creación del documento	2018-12-31	Leonardo Buitrago, Ricardo Ortíz, Camila Plata, Dairo Escobar.

## Acerca del SiB Colombia

**El SiB Colombia es la red nacional de datos abiertos sobre biodiversidad.** Esta iniciativa de país nace con el Decreto 1603 de 1994 como parte del proceso de creación del Sistema Nacional Ambiental (Sina), establecido en la Ley 99 de 1993, y es el nodo oficial del país en la Infraestructura Mundial de Información en Biodiversidad (GBIF). Su principal propósito es brindar acceso abierto a información sobre la diversidad biológica del país para la construcción de una sociedad sostenible. Además, facilita la publicación en línea de datos e información sobre biodiversidad, y promueve su uso por parte de una amplia variedad de audiencias, apoyando de forma oportuna y eficiente la gestión integral de la biodiversidad.

El SiB Colombia es una realidad gracias a la participación de cientos de organizaciones y personas que comparten datos e información bajo los principios de libre acceso, transparencia, cooperación, reconocimiento y responsabilidad compartida.

Lo coordina el Instituto Humboldt y es liderado por un Comité Directivo (CD-SiB), conformado por el Ministerio de Ambiente y Desarrollo Sostenible, los 5 institutos de investigación del SINA (Ideam, Invemar, IIAP, Sinchi e Instituto Humboldt), la Universidad Nacional de Colombia y Parques Nacionales Naturales de Colombia. El CD-SiB se apoya en un Comité Técnico (CT-SiB), grupos de trabajo para temas específicos y un Equipo Coordinador (EC-SiB) que cumple las funciones de secretaría técnica, acogiendo e implementando las recomendaciones del CD-SiB.

El SiB Colombia promueve la participación activa del gobierno, la academia, el sector productivo y la sociedad civil para lograr la consolidación de información confiable y oportuna que apoye la toma de decisiones a nivel nacional e internacional. Es además, el nodo oficial del país en la infraestructura mundial de información en biodiversidad -GBIF-.

La implementación del SiB Colombia, a partir del 2000, constituyó el primer resultado del nuevo enfoque de gestión de datos e información en el ámbito nacional y se encuentra articulado con el Sistema de Información Ambiental de Colombia (SIAC) como el subsistema de información que soporta el componente de biodiversidad.

# Contenidos

<b>Introducción</b>	<b>5</b>
<b>eBird</b>	<b>5</b>
<b>GBIF</b>	<b>6</b>
<b>iNaturalist</b>	<b>6</b>
<b>Xeno-canto</b>	<b>7</b>

## Introducción

La repatriación de datos consiste en integrar a los canales de participación del SiB Colombia los registros biológicos del país que han sido publicados por otros países, o redes internacionales de cooperación. Esta resulta una tarea vital para los nodos nacionales de biodiversidad en la construcción de las cifras de país y, en el caso de Colombia, un aporte fundamental en la consolidación del inventario nacional de la biodiversidad. Desde el Equipo Coordinador del SiB Colombia se han identificado cuatro (4) conjuntos de datos esenciales para ser repatriados: eBird, iNaturalist, GBIF y Xeno-canto. Cada una de estas fuentes de datos contienen información valiosa, tomada a partir de colecciones biológicas, iniciativas de ciencia ciudadana y registros con datos significativos asociados (ej. sonidos, imágenes). En conjunto, los datos disponibles para repatriación desde las fuentes mencionadas son 4'353.377 nuevos registros biológicos.

La repatriación de datos hacia los canales de participación del SiB Colombia [biodiversidad.co](http://biodiversidad.co) se desarrolla a través de los siguientes pasos:

1. Selección de las fuentes de datos
2. Solicitud y descarga de los datos
3. Estandarización *Darwin Core* y definición de los campos a publicar
4. Validación y limpieza de los datos
5. Elaboración de metadatos en IPT
6. Carga de datos y publicación en IPT
7. Integración a los canales de participación del SiB Colombia

Desde diferentes aproximaciones, cada uno de los conjuntos de datos que hacen parte de este proceso genera un aporte significativo en la consolidación de información sobre biodiversidad disponible a través del SiB Colombia:

## eBird

*eBird*<sup>1</sup> es el proyecto de ciencia ciudadana relacionado con biodiversidad más grande del mundo, con más de 100 millones de registros de aves contribuidos a diario por los pajareros alrededor del mundo. Una iniciativa colaborativa que cuenta con cientos de organizaciones socias, así como miles de expertos regionales, y cientos de miles de usuarios. eBird es liderado por el *Laboratorio de Ornitología de Cornell*.

Los datos de eBird documentan la distribución, abundancia, uso de hábitat y tendencias de las aves a través de listas de especies recolectadas dentro de un marco científico simple. Los pajareros indican cuándo, dónde y cómo fueron a pajarear, y luego llenan una lista de chequeo de todas las aves observadas o escuchadas durante la jornada.

---

<sup>1</sup> <https://ebird.org>

*eBird Colombia*<sup>2</sup> integra observaciones de las aves del país, gracias al esfuerzo de cientos de investigadores a nivel global que se suman a esta red de ciencia participativa nacional. Este conjunto de datos cuenta, a 31 de agosto 2018, con 3'239.923 registros de observación, 1.861 especies de aves correspondientes a 94 familias. La mayor cantidad de registros corresponde al orden de los Passeriformes. Las especies más observadas son *Coragyps atratus* (Chulo o Gallinazo) y *Tyrannus melancholicus* (Sirirí común) con más de 45 mil registros cada una.

Conjunto de datos repatriado disponible en: <https://doi.org/10.15472/zso3xd>

## GBIF

La *Infraestructura Global de Información sobre Biodiversidad - GBIF*<sup>3</sup>, es una red de gobiernos y organizaciones que facilita la publicación de datos y pone a disposición de todo el mundo, la información sobre diversidad biológica que investigadores y ciudadanos comparten en línea sin importar su ubicación geográfica. Colombia, considerado un país megadiverso, cuenta con una importante cantidad de datos disponibles a través de GBIF aportados por organizaciones de distintos países, que cuentan con evidencias de la presencia de especies en el territorio nacional.

Este conjunto de datos comprende 1'082.453 registros biológicos publicados para Colombia a través de la GBIF (a 28 de diciembre de 2018). Estos datos han sido aportados por 324 organizaciones alrededor del mundo, que han recopilado datos sobre la diversidad biótica del país y permiten complementar el inventario nacional de la biodiversidad que se ha venido consolidando a través del SiB Colombia. Aproximadamente el 85% de los datos corresponden a especímenes preservados en colecciones biológicas. Se encuentran allí registros biológicos tomados desde el siglo XVII hasta la actualidad.

Conjunto de datos repatriado disponible en: <https://doi.org/10.15472/xqv1re>

## iNaturalist

Es una de las aplicaciones de ciencia participativa más populares del mundo, *iNaturalist*<sup>4</sup> ayuda a identificar las plantas y los animales que nos rodean. Esta iniciativa permite conectarse con una comunidad de más de 750.000 científicos y naturalistas que pueden ayudar a los usuarios a aprender más sobre la naturaleza. Además, al registrar y compartir las observaciones, se crean datos con calidad de investigación para que los científicos puedan conocer y proteger la naturaleza. *iNaturalist* es una iniciativa conjunta de la *Academia de Ciencias de California* y la *National Geographic Society*. En Colombia, esta apuesta tecnológica que articula ciencia y conocimiento no experto, recibe el nombre de **Naturalista Colombia**<sup>5</sup> y es una herramienta esencial dentro de la estrategia nacional de ciencia participativa que lidera el Instituto Humboldt.

---

<sup>2</sup> <https://ebird.org/colombia>

<sup>3</sup> <https://www.gbif.org>

<sup>4</sup> <https://www.inaturalist.org>

<sup>5</sup> <https://colombia.inaturalist.org>

Los datos que reciben confirmación de identificación por parte de tres o más usuarios de la plataforma son elevados a 'Grado de investigación' (Research grade), esto permite a iNaturalist mantener un mecanismo permanente para segmentar estos datos. Los registros en esta categoría para Colombia (que son los que hacen parte de la repatriación) suman 26.967, con corte a 28 de diciembre 2018. Más de 4.000 especies son aportados a partir de este conjunto de datos, en su gran mayoría evidencia de aves, plantas e insectos.

Conjunto de datos repatriado disponible en: <https://doi.org/10.15472/gsxqe8>

## Xeno-canto

Xeno-canto es una iniciativa web<sup>6</sup> dedicada a compartir sonidos de aves de todo el mundo. Científicos investigadores, observadores de aves o simplemente usuarios curiosos pueden escuchar, descargar y explorar las grabaciones de los sonidos de aves presentes en toda esta colección. La *Fundación Xeno-canto para Sonidos de la Naturaleza* (Xeno-canto, XC) tiene como objetivo popularizar la grabación, mejorar la accesibilidad y aumentar el conocimiento de sonidos de aves en todo el mundo. La fundación tiene su sede en Holanda, fue fundada en 2008 y es conocida oficialmente por su nombre holandés "Stichting Xeno-canto voor Natuurgeluiden". Las grabaciones presentes en Xeno-canto, son compartidas por una creciente comunidad de miles de grabadores de todo el mundo, observadores de aves aficionados y profesionales por igual. El objetivo de Xeno-canto es tener representación de todos los sonidos de las aves del mundo, es decir, de todos los taxones, inclusive a nivel de subespecie, su repertorio completo, toda la variabilidad geográfica, en todas las etapas de desarrollo.

El conjunto de datos de Xeno-canto comprende 4.034 registros sonoros de más de 780 especies de aves de Colombia. Cuenta con datos de 1989 al año 2017, e incluye más de 300 aves identificadas a nivel de subespecie.

Conjunto de datos repatriado disponible en: <https://doi.org/10.15472/lo4iij>

---

<sup>6</sup> <https://www.xeno-canto.org>