

2026-10-19

Datos abiertos de investigación

Ciencia Abierta PEDECIBA (Open Science Uruguay 2026)

Florencia Grattarola



Objetivos

Datos abiertos de investigación

❑ QUÉ

Introducir conceptos sobre los datos abiertos de investigación.

Ver un caso de ejemplo.

❑ POR QUÉ

Entender por qué son importantes la gestión y la apertura de los datos de investigación.

QUÉ

Conceptos

QUÉ
Conceptos

¿Qué son los datos abiertos de investigación?

Datos de investigación

- Son la **evidencia** comúnmente aceptada por la comunidad académica específica como necesaria para apoyar científica y técnicamente la investigación.
- Se generan como parte de **actividades de investigación** como experimentos, mediciones, encuestas, entrevistas, observaciones, etc.
- Y son tan variados como las disciplinas académicas y los métodos con que se generan



Datos abiertos

- El conocimiento es abierto si cualquiera es libre de **acceder** a él, **usarlo, modificarlo y compartirlo**, estando sujeto a lo sumo a medidas que preserven su autoría y su apertura.

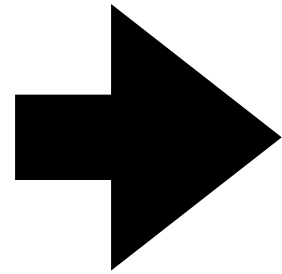


Open
Definition

<https://opendefinition.org/od/2.1/es/>

Requisitos	
Licencia abierta	Deben estar en el dominio público o disponible bajo una licencia abierta
Acceso	Deben estar disponible como un todo y deberían ser descargable de manera gratuita a través de Internet
Legible por máquina	Deben proporcionarse de manera que sean procesables por una computadora
Formato abierto	Deben proporcionarse en un formato abierto (que no impone restricciones)

	man	Forest	grass
1988	65	28	7
1998	52	20	28
2008	59	33	7
2018	75	22	2



CSV

year	man	forest	grass
1988	65	28	7
1998	52	20	28
2008	59	33	7
2018	75	22	2



Term	Definition
man	Percentage of land that is managed
forest	Percentage of land that is forested with trees
grass	Percentage of land that is grassland
year	Year in which the measure was extracted

QUÉ
Conceptos

Principios FAIR y CARE

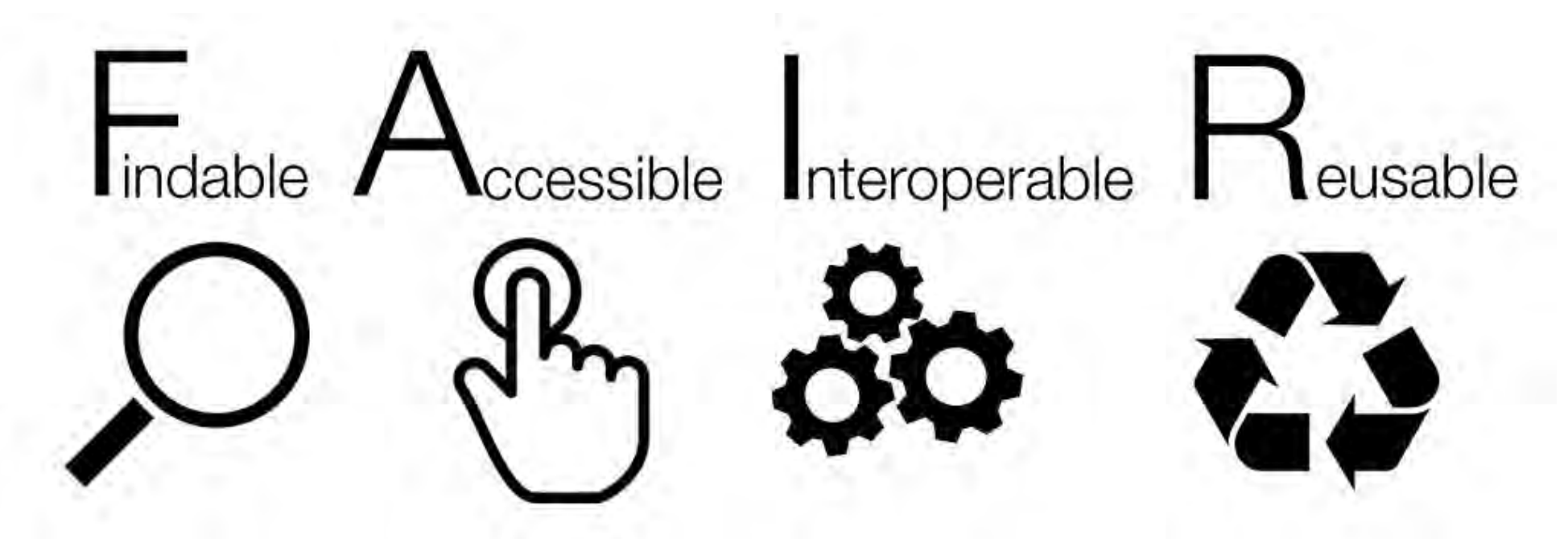
Principios de gestión de datos

“Lo más abierto posible, tan cerrado como sea necesario”

... siempre que fuera ética, legal y comercialmente posible

Principios FAIR para que los datos sean:

- **Encontrables**
- **Accesibles**
- **Interoperables**
- **Reutilizables**



<https://www.go-fair.org/fair-principles/>

Principios de gestión de datos

Ser justos (FAIR) hacia la creación de formas apropiadas para la gobernanza de datos indígenas

Principios CARE:

- Beneficio colectivo
- Autoridad para controlar
- Responsabilidad
- Ética



<https://www.gida-global.org/careprinciples>

QUÉ
Conceptos

Gestión de datos: planes de gestión de datos

Gestión de datos

- Permite un manejo efectivo y responsable de la investigación.
- Incluye la recolección, documentación, almacenamiento, intercambio y preservación a largo plazo de los datos de investigación dentro y más allá del curso de un proyecto determinado.



Datos
Abiertos

abierto $\neq$ accesible

Externo
(para beneficio comunitario)

Datos FAIR

Interno
(para uso propio)

Datos gestionados

Plan de Gestión de Datos

- Es un documento formal que detalla las estrategias que serán usadas para la **gestión eficiente de los datos generados o recolectados** durante el proyecto de investigación y luego de su cierre.
- Es una herramienta de apoyo en cada etapa del ciclo de vida de los datos: **recolección, generación, tratamiento y publicación.**



Componentes

**Datos generados
o recopilados y
documentación**

**Almacenamiento y
preservación a
largo plazo**

**Intercambio y
reutilización
los datos**

**Cumplimiento
ético y legal**



Documentación

Ejemplo a nivel de proyecto y de los datos



<https://assets.britishecologicalsociety.org/2026/09/21103535/BES-Data-management-guide-2026.pdf>

Project-level	Data-level
• The project aim, objectives and/or hypotheses, or research question. • Personnel involved throughout the project, including who to contact with questions. • Details of sponsors. • Data collection methods, including details of instrumentation and environmental conditions during collection, copies of collection instructions if applicable. • Data standards used. • Data structure and organisation of files: files description and format encoding, relationships between files. • Temporal, spatial and/or taxonomical coverage. • Data completeness and known gaps. • Software used to prepare and read the data. • Procedures used for data processing, including quality control and versioning and the dates these were carried out. • Known problems that may limit accessibility and data use. • Data validation process. • Instructions on how to cite the data. • Intellectual property and other licensing considerations.	• Names, labels and descriptions for variables. • Detailed explanation of codes used. • Definitions of acronyms or specialist terminology. • Reasons for missing values and missing value conventions. • Units of measurement (critical for interpretation). • Measurement precision and uncertainty (e.g. body mass to nearest gram, GPS accuracy ± 50 cm). • Derived data created from the raw file, including the code or algorithm used to create them.

QUÉ
Conceptos

Estándares de datos

Estándares de datos

- Cuando estandarizamos los datos, los ponemos en un lenguaje común que otrxs (incluidas las máquinas) pueden entender.
- Un ejemplo clásico es cómo documentar las **fechas**.
 - Por ejemplo: 12-01-2020 ¿es 12 de enero o 1ero de diciembre?
- El estándar más conocido para los datos de biodiversidad es **Darwin Core**.



<https://www.tdwg.org>



<https://dwc.tdwg.org/terms/>

Calidad de datos



<https://assets.britishecologicalsociety.org/2026/09/21103535/BES-Data-management-guide-2026.pdf>

Figure 2: **Examples of checks that could be carried out during quality assurance.**

id	latitude	longitude	date	start_time	end_time	total_flowers	vegetation_height	cloud_cover
1	50.954718	-0.62347412	04/05/2026	10:06	6:20	NA	50	All or mostly blue
2	50.954718	-0.62347412	2026-05-31	9:15	15:30	0	10	All or mostly cloud
2	50.954718	-0.62347412	2026-06-04	9:15	15:30	0	10	All or mostly cloud
3	50.954718	-0.62347412	2026-06-31	9:30	10:50	2200	85	Half blue and half cloud
4	50.594718	-0.62347412	2026-07-04	NA	16:00	240	250	50% cloud

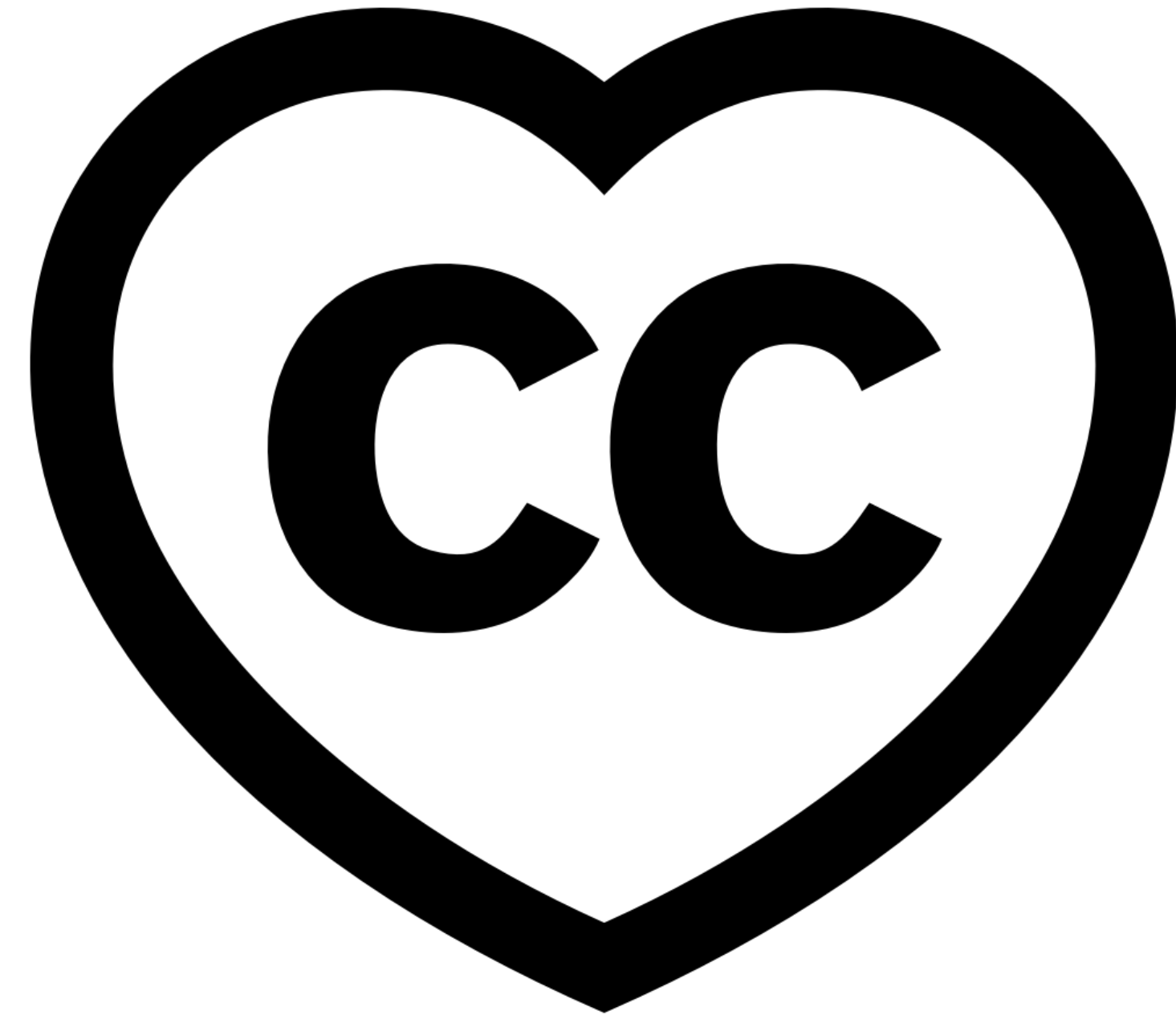
Data structure and methodology based on the UK Pollinator Monitoring Scheme
<https://doi.org/10.5285/5841bfae-471b-498c-af89-c30e5140d6c5>

QUÉ
Conceptos

Licencias: uso de datos y atribución

Licencias

- Las licencias ofrecen una forma sencilla y estandarizada de otorgar al público permiso para compartir y utilizar tu obra creativa bajo las condiciones que elijas.
- No son automáticas. **Tenés que declarar la licencia que hayas elegido.**



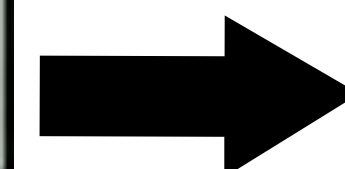
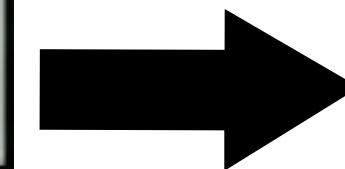
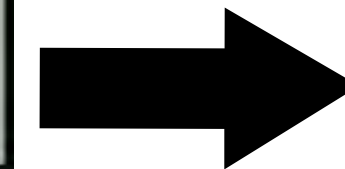
En dominio público <3



La gran ola de Kanagawa por Katsushika Hokusai

En Uruguay <https://dominiopublico.uy/>

Pero si tuviera otra licencia



Licencias para datos abiertos

	Licencia		Aplicación	URL
	CCZero	CC0	contenido, datos	https://opendefinition.org/licenses/cc-zero
	Atribución	CC-BY	contenido, datos	https://opendefinition.org/licenses/odc-pddl
	Atribución Compartir Igual	CC-BY-SA	contenido, datos	https://opendefinition.org/licenses/cc-by
 Open Data Commons	Public Domain Dedication and License	PDDL	datos	https://opendefinition.org/licenses/odc-by
	Atribución	ODC-By	datos	https://opendefinition.org/licenses/cc-by-sa
	Open Database License	ODbL	datos	https://opendefinition.org/licenses/odc-odbl

Para software



<https://opensource.org/licenses/>

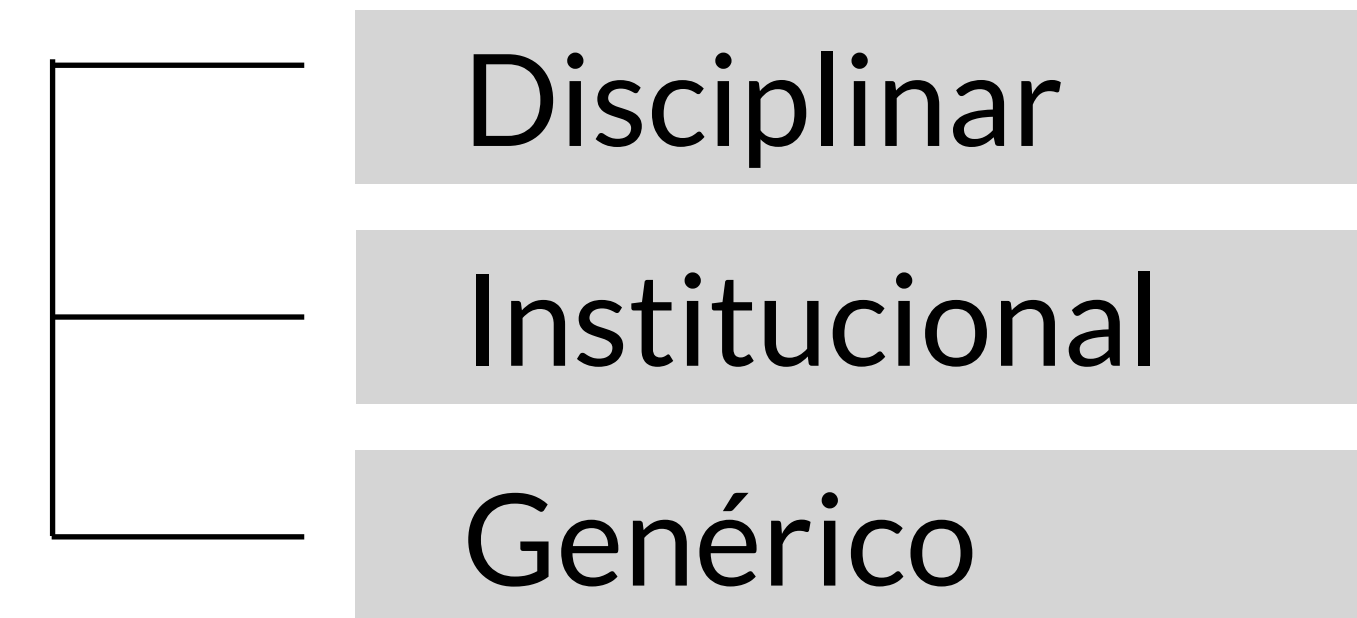
QUÉ

Conceptos

**Repositorios: institucionales,
genéricos y de disciplina**

Repositorios de datos

- Una pregunta importante sobre el intercambio de datos es: ¿dónde debo depositar mis datos?
- ¡Hay muchas opciones de repositorios! Hay que elegir con cuidado :)



Lo que no deberíamos hacer es compartir datos como archivos suplementarios :/

Supporting Information	
Filename	Description
supp-file.docx	S1. The collected data used for the study
Please note: The publisher is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing content) should be directed to the corresponding author for the article.	

Repositorios de datos

- **Disciplinar:** repositorio especializado para conservar tus datos de acuerdo con los estándares reconocidos en tu disciplina.
- **GBIF** es el principal repositorio de datos primarios sobre biodiversidad. También es un agregador de datos, lo que significa que los datos de otras bases de datos terminan en GBIF.



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



Genbank

<https://www.ncbi.nlm.nih.gov/genbank/>

Repositorios de datos

- **Institucional:** repositorio de tu institución, universidad u organismo de financiamiento.
- Por lo general, se trata de repositorios genéricos (no específicos de una disciplina).



Repositorios de datos

- **Genérico:** repositorio público (cualquier persona con un usuario puede subir datos).
- Almacenan una gran variedad de tipos de datos procedentes de diversas disciplinas.
- Una desventaja es que, por lo general, no se controla la calidad de los datos ni de los metadatos.



zenodo

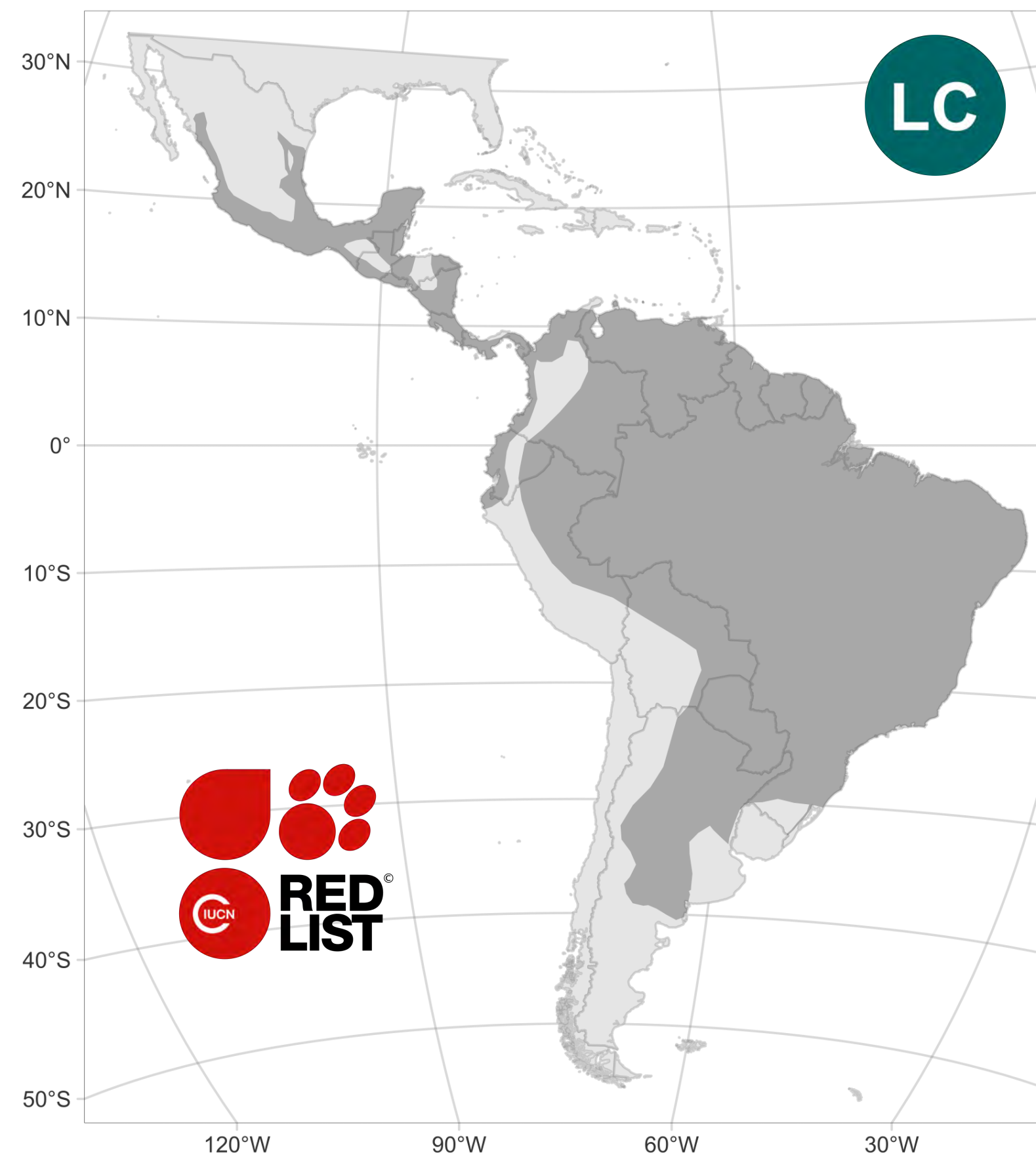


QUÉ

Un caso de ejemplo

QUÉ

Un caso de ejemplo



IUCN (2022)

Una nueva especie en Uruguay: el yaguarundi

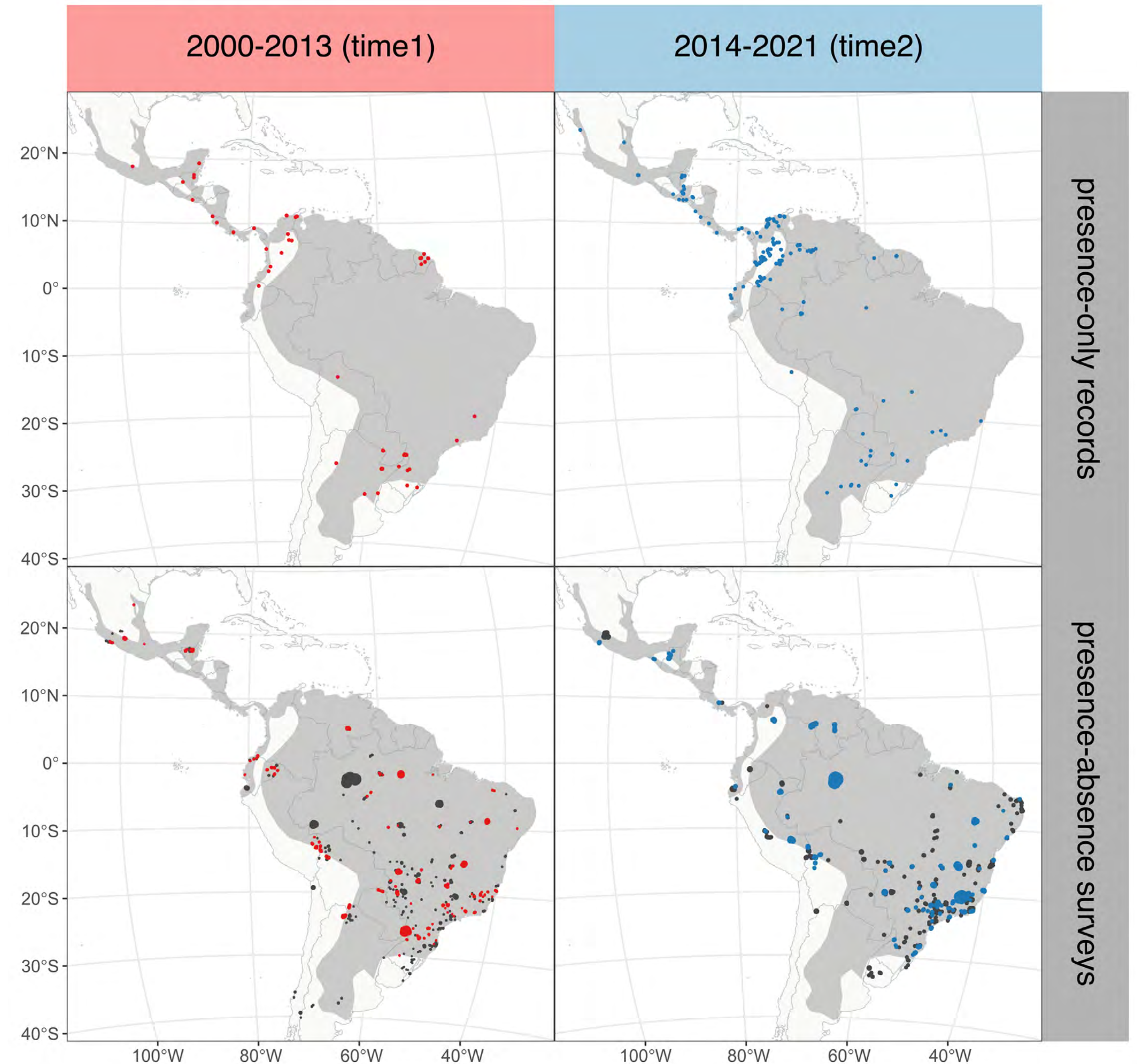


JULANA (2015)

QUÉ

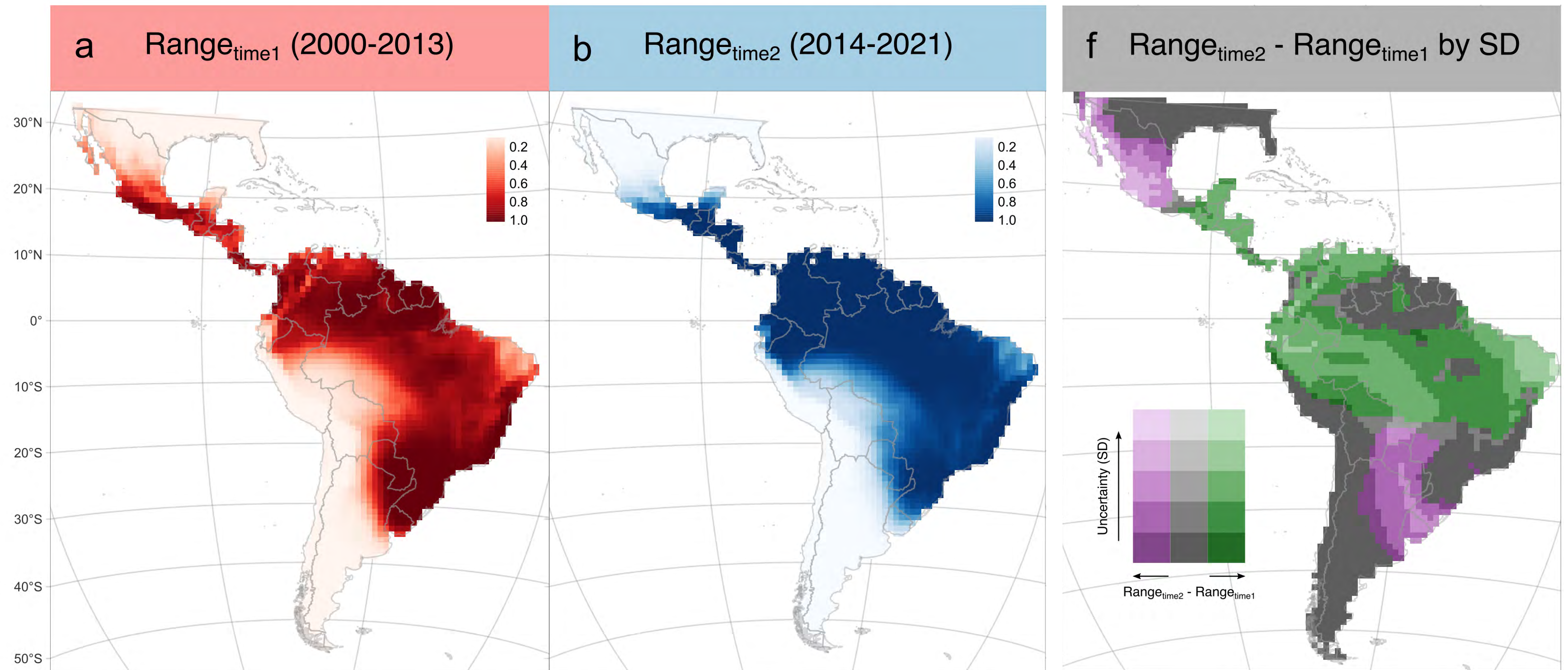
Un caso de ejemplo

Compilación de datos de presencia y presencia-ausencia para estudiar cambios temporales haciendo uso de modelos bayesianos (integración de datos)

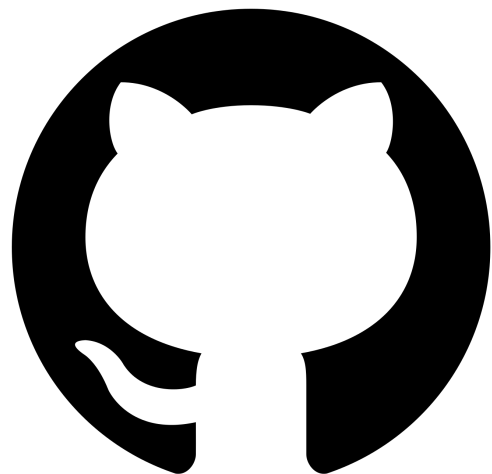


QUÉ

Un caso de ejemplo



repositorio



main		
Go to file		
Code		
bienflorenzia Update README.md 7177b1f · 2 years ago 50 Commits		
code	add records per country	2 years ago
data	Update carnivores.csv	2 years ago
metadata	update bib and csv references' files	2 years ago
.gitignore	Update .gitignore	2 years ago
LICENSE	create repo	2 years ago
MIAU.Rproj	Rename Neotropical-carnivores-data...	2 years ago
README.md	Update README.md	2 years ago
_quarto.yml	Update _quarto.yml	2 years ago

https://github.com/bienflorenzia/MIAU

archivo del repositorio



Published December 4, 2024 | Version v1.0

bienflorenzia/MIAU

Grattarola, Florencia · Tschernostárova, Katerina · Keil, Petr

Software Open

Show affiliations

MIAU: An analysis-ready database on presence-only and presence-absence data of Neotropical carnivores (Mammalia: Carnivora) from 2000 to 2021

In the last decade, databases of records of species observed at the same location at different points in time over large spatial extents have been made available. Unfortunately, these sources are scarce in regions such as Latin America. We present a dataset of 60,179 point occurrences (i.e. presence-only data, PO) and 45,468 camera trap survey records (i.e. presence-absence data, PA) for 63 species of carnivores of the Neotropical region from 2000 to 2021. We collated the data from various sources, including 64 newly digitised bibliographic references. We cleaned, taxonomically harmonised, and standardised the data following the Darwin Core and Humboldt Core standards and present them here as csv files. We have also made these data fit for analyses by aggregating the data into two time periods (time1: 2000-2013 and time2: 2014-2021), with PO grid cell counts of 100 x 100 km, and PA polygons of varying size, presented as geopackage files. These data can be used for large-scale species distribution models, calculation of population trends, extinction risk analyses, and educational purposes.

https://doi.org/10.5281/zenodo.14278694

data paper

Nature Conservation 58: 11–30 (2025)
DOI: 10.3897/natureconservation.58.140644

Data Paper

MIAU: An analysis-ready dataset on presence-only and presence-absence data of Neotropical carnivores (Mammalia, Carnivora) from 2000 to 2021

https://doi.org/10.3897/natureconservation.58.140644

artículos de investigación

RESEARCH ARTICLE

Journal of Biogeography WILEY

Integrating presence-only and presence-absence data to model changes in species geographic ranges: An example in the Neotropics

Florencia Grattarola¹ · Diana E. Bowler^{2,3,4} | Petr Keil¹

https://doi.org/10.1111/jbi.14622

Contents lists available at ScienceDirect

Global Ecology and Conservation

journal homepage: www.elsevier.com/locate/gecco

A continental-wide decline of occupancy and diversity in five Neotropical carnivores

https://doi.org/10.1016/j.gecco.2024.e03226

data paper

artículo de investigación



Nature Conservation 58: 11–30 (2025)
DOI: 10.3897/natureconservation.58.140644

Data Paper

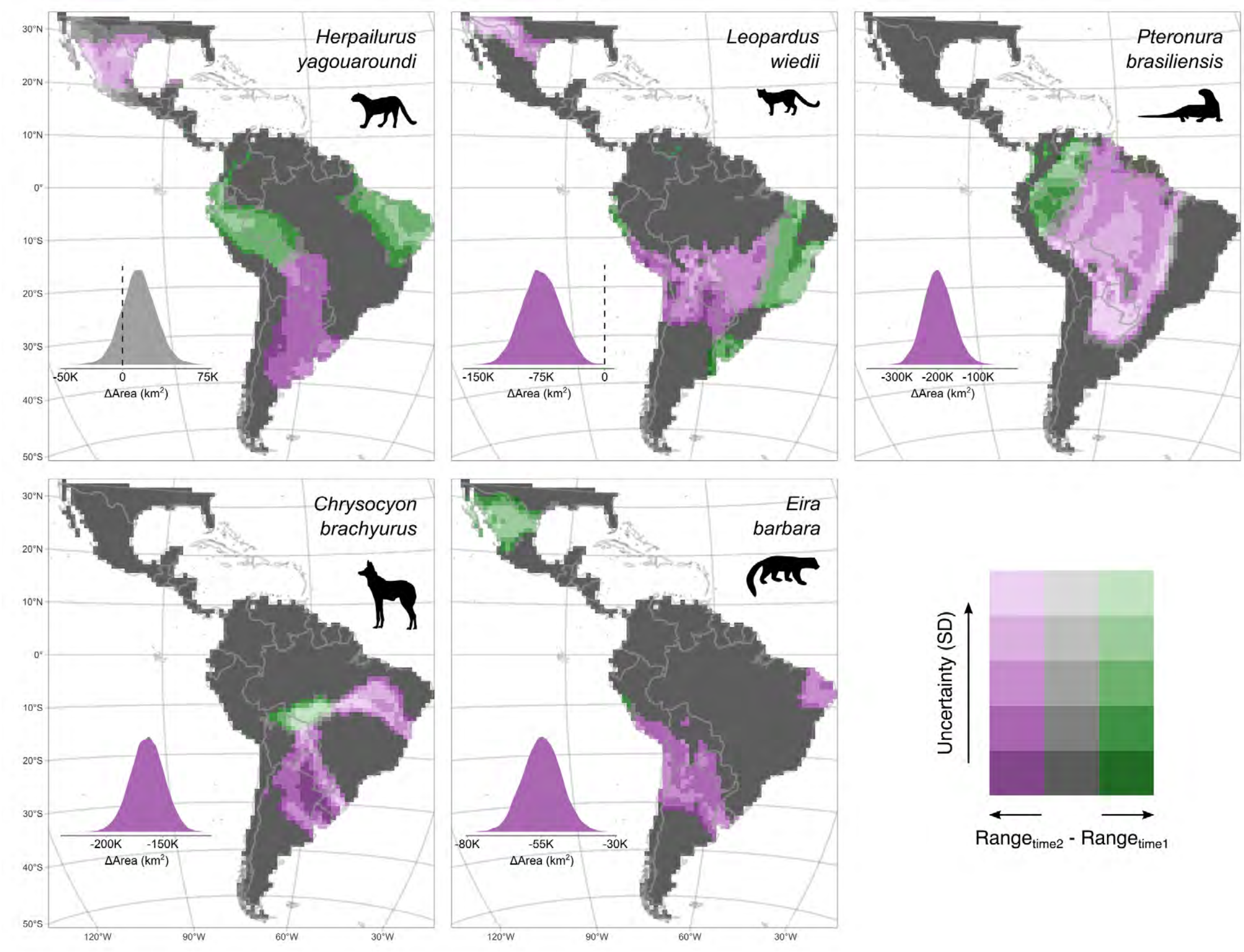
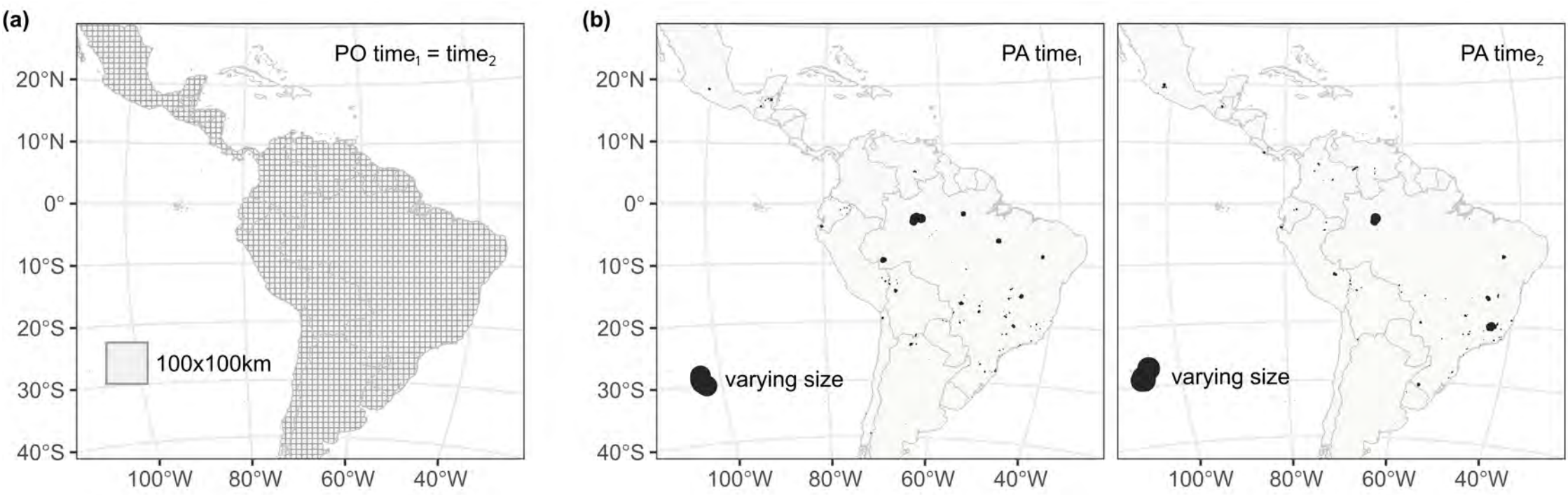
MIAU: An analysis-ready dataset on presence-only and presence-absence data of Neotropical carnivores (Mammalia, Carnivora) from 2000 to 2021



Contents lists available at [ScienceDirect](#)
Global Ecology and Conservation
journal homepage: www.elsevier.com/locate/gecco



A continental-wide decline of occupancy and diversity in five Neotropical carnivores

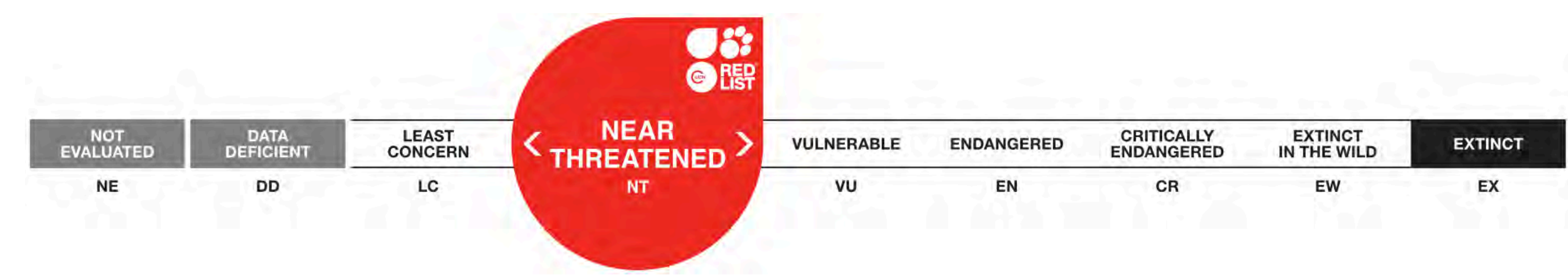


Yaguarundi

Herpailurus yagouaroundi

Recent evidence suggests that, between a timeframe spanning from 2000–2013 and 2014–2021, the geographic range of the Jaguarundi has retreated in the central part of Argentina, Paraguay, and Uruguay (Grattarola et al. 2023). The extent of reduction in its extent of occurrence (EOO) is unknown.

IUCN (2026)



POR QUÉ

Importancia

POR QUÉ

Importancia

Propósito de la ciencia

¿Por qué hacemos ciencia?

- Somos las personas responsables, designadas por la sociedad, de **ampliar el conocimiento disponible** sobre el mundo que habitamos
- **No trabajamos para nosotrxs mismxs**, si no para quienes sostienen nuestro trabajo.



Mariska Gewald

POR QUÉ

Importancia

"Todo es un remix"

Kirby Ferguson. Everything is a Remix (www.everythingisaremix.info)

Los *bienes comunes* del conocimiento

- La ciencia es un **proceso acumulativo** que se construye en base a conocimientos previamente descubiertos.
- Podemos hacer esto porque el conocimiento se ha hecho accesible y podemos acceder a él (“**fondo común**”).

Elinor Ostrom. Governing the Commons: The Evolution of Institutions for Collective Action



Mariska Gewald

POR QUÉ

Importancia

Reforma de la evaluación e incentivos

Incentivos en ciencia abierta

No debe quedar sólo en manos de investigadorxs.

Iniciativas arriba hacia abajo (*top-down*):

- Reformar el **sistema de evaluación científica** (investigación, investigadorxs e instituciones de investigación).
- Implementar esquemas de incentivos a la ciencia abierta (**participativa**).
- **Capacitar** en ciencia abierta y apoyar.



Mariska Gewald

¿Y AHORA?

Primeros pasos

¿Cómo arranco?

1. **Empezá de a poco:** create un ORCID iD (¡usalo!)
2. **Arrancá temprano:** gestioná tus datos desde el comienzo (no tiene que ser perfecto)
3. **Involucrate:** sumate a la comunidad :)

Comunidades de Ciencia Abierta



¡Gracias!

Esta presentación se puede descargar y reutilizar

Créditos: Autoría y fuente incluidas en el pie de cada recurso/imagen



Czech University
of Life Sciences Prague

 flograttarola.com |  [@flograttarola.bsky.social](https://flograttarola.bsky.social) |  ecoevo.social/@flograttarola