Biodiversidata: A Collaborative Initiative Towards Open Data Availability in Uruguay

Florencia Grattarola | University of Lincoln

flograttarola.com | @flograttarola
Uruguay is deeply ignored in terms of biodiversity.
What is the scenario?
What is the scenario?

Uruguay is one of the countries of Latin America with the lowest levels of access to biodiversity data.

The public sources of biodiversity data in the country are not digital.

GitHub: https://github.com/bienflorencia/RBiodiversidata/tree/master/GBIF%20Latin%20America
People are willing to share their data

- 86% Yes
- 14% No

OBSTACLES
1. Time and effort to make data available
2. Absence of data management plans
3. Lack of recognition for their effort

INCENTIVES/MOTIVATIONS
1. Create networks with other scientists
2. Get recognition if their data are used by others

What is the scenario?
What did we do?
What did we do?

PLAN

• Create a network of experts sharing data and knowledge.
• Simplify data processing workflow.
• Get credit for data and research.
What did we do?

GOALS

• Collect the maximum possible amount of primary data from vertebrate, invertebrate and plant species.
• Use it to collaborative generate global impact scientific research.
• Make the data available free and open.
What did we do?

**Data Collection**

- **taxa**
- **geographic location**
- **date**

<table>
<thead>
<tr>
<th>ID</th>
<th>archive</th>
<th>ID camera</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Specie</th>
<th>X</th>
<th>Y</th>
<th>Localidad</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-2.12327</td>
<td>-53.7549</td>
<td>Hominoides</td>
<td>San José</td>
<td>Casa de Edifa</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-2.12327</td>
<td>-53.7549</td>
<td>Mazama gouazoubira</td>
<td>Canelones</td>
<td>Casa de Edifa</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-2.12327</td>
<td>-53.7549</td>
<td>Mazama gouazoubira</td>
<td>Canelones</td>
<td>Casa de Edifa</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-2.12327</td>
<td>-53.7549</td>
<td>Mazama gouazoubira</td>
<td>Canelones</td>
<td>Casa de Edifa</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-2.12327</td>
<td>-53.7549</td>
<td>Mazama gouazoubira</td>
<td>Canelones</td>
<td>Casa de Edifa</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>-2.12327</td>
<td>-53.7549</td>
<td>Mazama gouazoubira</td>
<td>Canelones</td>
<td>Casa de Edifa</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>-2.12327</td>
<td>-53.7549</td>
<td>Mazama gouazoubira</td>
<td>Canelones</td>
<td>Casa de Edifa</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>-2.12327</td>
<td>-53.7549</td>
<td>Mazama gouazoubira</td>
<td>Canelones</td>
<td>Casa de Edifa</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>-2.12327</td>
<td>-53.7549</td>
<td>Mazama gouazoubira</td>
<td>Canelones</td>
<td>Casa de Edifa</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>-2.12327</td>
<td>-53.7549</td>
<td>Mazama gouazoubira</td>
<td>Canelones</td>
<td>Casa de Edifa</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**GBIF**

+ [Open Access Data Portal](https://www.gbif.org/)
What did we do?

Data Collection

<table>
<thead>
<tr>
<th>taxa</th>
<th>geographic location</th>
<th>date</th>
</tr>
</thead>
</table>

(sources)

FERRUGINOUS PYGMY OWL *Glaucidium brasilianum*

Two photographed at Arroyo de las Pájaras, c.38 km west of Treinta y Tres City. Apto, Treinta y Tres, on 2 March 2007 (A. Rocchi & G. Mondón). One photographed and another heard at Quebrada de los Cuervos, dpto. Treinta y Tres, on 3 March 2009 (CC). The species was also observed there in 1999 (CC). Found near Paso Pávano, río Yaguarón, eastern dpto. Cerro Largo on 15 (one) and 16 (two) January 2011; one bird was sound-recorded (ABA, JLM, XC9301). Also photographed at three sites south of Paso Centurión on 6-7 March (DP, JLM, JSA), 9 March (ABA) and 1 November 2011 (JLM, JSA, D. Gil). Singles and pairs were seen repeatedly and photographed at Paso Averías, río Cebollati, southern dpto. Treinta y Tres, on 19-24 April 2011 (AR & M. Abreu). There are very few previous reports for Uruguay. Wethington (1926) collected a female near Lascano, dpto. Rocha in February 1921, while Tremolada (1927) took two at Arroyo Grande, dpto. Flores in 1891 and reported another in dpto. Lavalleja taken prior to 1927, and one was collected at Paso de las Piedras, río Negro, dpto. Dánszko, in April 1961 (Cuollo & Gerzstein 1962), but ours are the first records since then. Probably not as rare as previously considered (especially around Centurión) and the species appears to be widespread over the country. It is scarce in Entre Ríos, Argentina (de la Peña 1997) and rare in Río Grande do Sul, Brazil (Belon 1984).

**ESPECIES DE ANFIBIOS Y REPTILES REGISTRADAS EN PASO CENTURIÓN - RÍO YAGUARÓN**

<table>
<thead>
<tr>
<th>Especies</th>
<th>Nombre común</th>
<th>Sitio I</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lepidodactyliidae</td>
<td>Rana pidioca (X)</td>
<td></td>
</tr>
<tr>
<td>Lepidodactylus nitidus</td>
<td>Rana saltadora (X)</td>
<td></td>
</tr>
<tr>
<td>Pseudopaludicola fuliginosa</td>
<td>Rana macacucho (X)</td>
<td></td>
</tr>
<tr>
<td>Pseudodactylus minutus</td>
<td>Rana boydiana (X)</td>
<td></td>
</tr>
<tr>
<td>Hylidae</td>
<td>Rana tropidoca (X)</td>
<td></td>
</tr>
<tr>
<td>Hyla minutula</td>
<td>Rana nyctea (X)</td>
<td></td>
</tr>
</tbody>
</table>
What did we do?

Data Collection

- taxa
- geographic location
- date

(sources)

GBIF
**What did we do?**

**Cleaning and Standardisation (semi-automated)**

- **OpenSci**
- **ITIS**
- **GeoNames**

<table>
<thead>
<tr>
<th>OLD Species</th>
<th>NEW from The Reptile Database</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anops kingi</td>
<td>Anphisabraea kingi</td>
</tr>
<tr>
<td>Lepidosaurus murneal</td>
<td>Epicte murneal</td>
</tr>
<tr>
<td>Lophis jaegeri</td>
<td>Erythrolepis jaegeri</td>
</tr>
<tr>
<td>Lophis poeoiogryna</td>
<td>Erythrolepis poeoidgryna</td>
</tr>
<tr>
<td>Lophis almadensis</td>
<td>Erythrolepis almadensis</td>
</tr>
<tr>
<td>Lophis mitrans</td>
<td>Erythrolepis mitrans</td>
</tr>
<tr>
<td>Chelidophorus ocellifer</td>
<td>Anneluella ocellifera</td>
</tr>
</tbody>
</table>

- occurrenceID
- scientificName
- scientificNameAuthorship
- vernacularName
- kingdom
- phylum
- class
- order
- family
- genus
- specificEpithet
- infraspecificEpithet
- countryCode
- stateProvince
- verbatimLocality
- decimalLatitude
- decimalLongitude
- georeferenceSources
- georeferencedBy
- eventDate
- year
- month
- day
- basisOfRecord
- institutionCode
- collectionCode
- catalogNumber
- recordedBy
- recordNumber
- identifiedBy
- associatedReferences

**Biodiversity Information Standards**

**taxa geographic location date**
What did we do?

69,380 occurrence records (non-duplicated)
673 species
Data Paper

What did we do?

Biodiversidata: An Open-Access Biodiversity Database for Uruguay


https://doi.org/10.3897/BDJ.7.e36226

Repository

Biodiversidata: An Open-Access Biodiversity Database for Uruguay

June 20, 2019

https://doi.org/10.5281/zenodo.2650169

Scripts

rBiodiversidata

Welcome to Biodiversidata’s R repository

The Uruguayan Consortium of Biodiversity Data is a collaborative association of entities that aim to make the biodiversity data of Uruguay accessible to researchers, including databases, publications, maps, reports, and infographics. This repository contains scripts and data used in their project.

You can find here scripts about:
- (Paper in prep): Analysis of Tetrapods’ Hotspots
- Data Cleaning & Standardisation
- Data Analysis about GBIF records of Latin America
- Directory of useful files

https://github.com/bienflorencia/rBiodiversidata
What did we do?

Research Paper
What did we do?
What did we learn?
What did we learn?

There is a major gap between being willing to share and effectively doing so.

21.5%

- Contacted: 65
- Replied Positively: 53
- Sent Data: 14
What did we learn?

There is insufficient support for the maintenance of national collections and digitisation of specimens.

Given lack of support, people invest personal time (sometimes money). They don’t feel this is valued, so who ‘owns’ the data gets fuzzy.
Data-sharing is perceived as a negative subject given past experiences

Public funded data should be open. Yes, but…

It is important to understand what is behind non-sharing practices. To make it compulsory may only make people more reluctant.
What did we learn?

- Inexistent culture of standardised data-sharing among researchers
- The lack of data management plans can be a greater barrier even for those willing to share.
- There is a rooted habit of asymmetric use of the data generated, there is no vision of reutilisation of data.
There is a lack of strategic plans towards open science in most of Uruguay’s research institutions.

Total absence of incentive structures (not incorporated in research assessments), making data-sharing a personal decision more than an institutional one.
What did we learn?

• **The path is long, full of ups and downs.** In low-resource research scenarios imposition does not help. So, we need to be patience and keep going. But with each step, push a little bit further.

• **Target those willing.** Directing the project to individual researchers/experts and not institutions was key in reducing time. But we need to start engaging more people.

• **Centralising the cleaning and standardisation processes.** This allowed researchers to send their raw records and save time. At the end, it enlarged the amount of data being collated. But we need to start training others on learning the language of data-sharing.

• **Communication the results and outcomes.** This made the project gain visibility. But we still need to engage the citizens and involve them in the conversation.
A socially committed science must be open, must enable participation and seek knowledge democratisation.
Building a global infrastructure for biodiversity data. Together.
22 - 25 October
Leiden - The Netherlands

Better data - better science - better policies