SICol

SICol stands for Information System on Collections of Biotechnological Interest

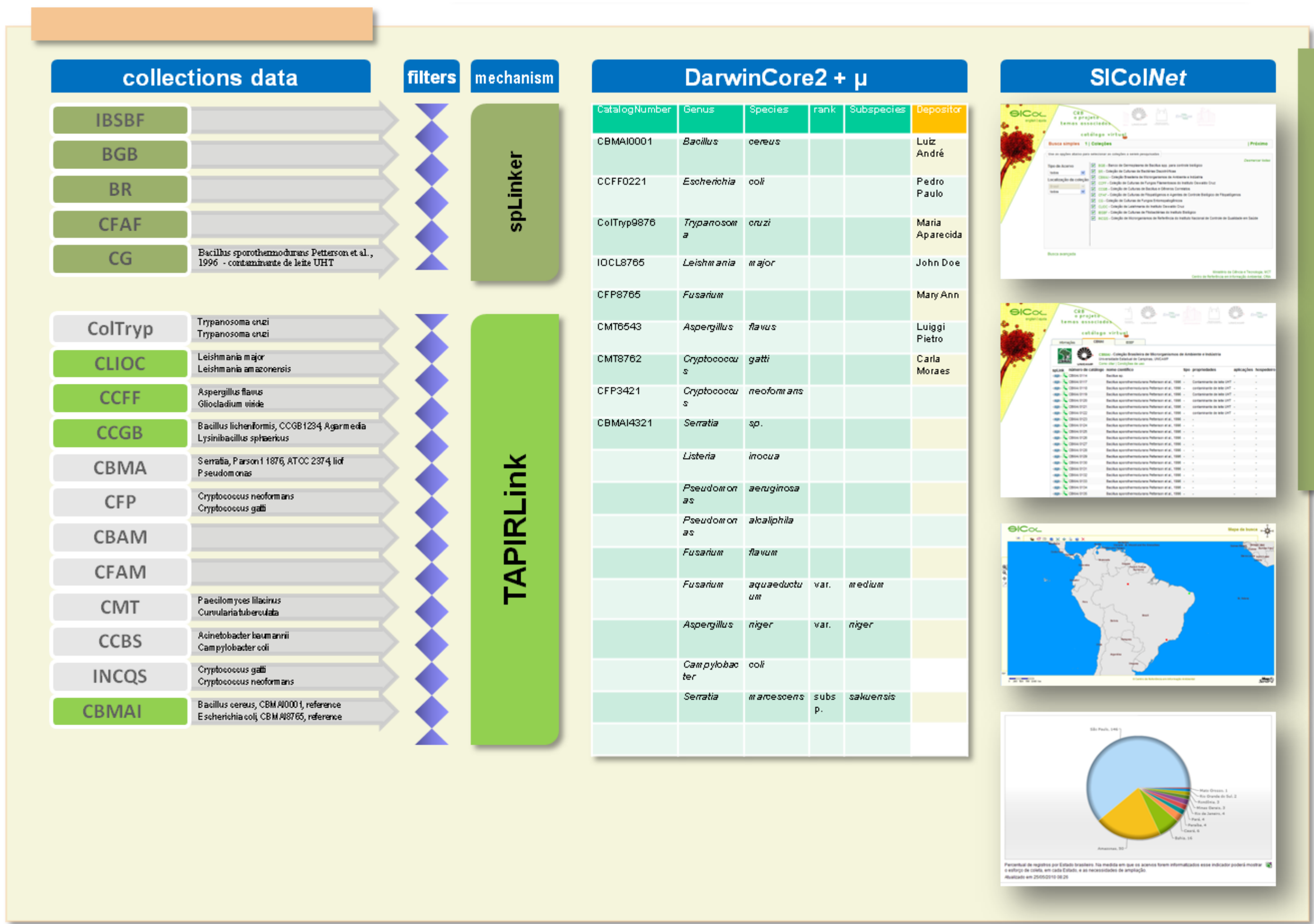

Alexandre Marino, Diego Rocha, Sidnei de Souza, Dora Ann Lange Canhos, Vanderlei Perez Canhos

Centro de Referência em Informação Ambiental, CRIA, Brazil, www.cria.org.br

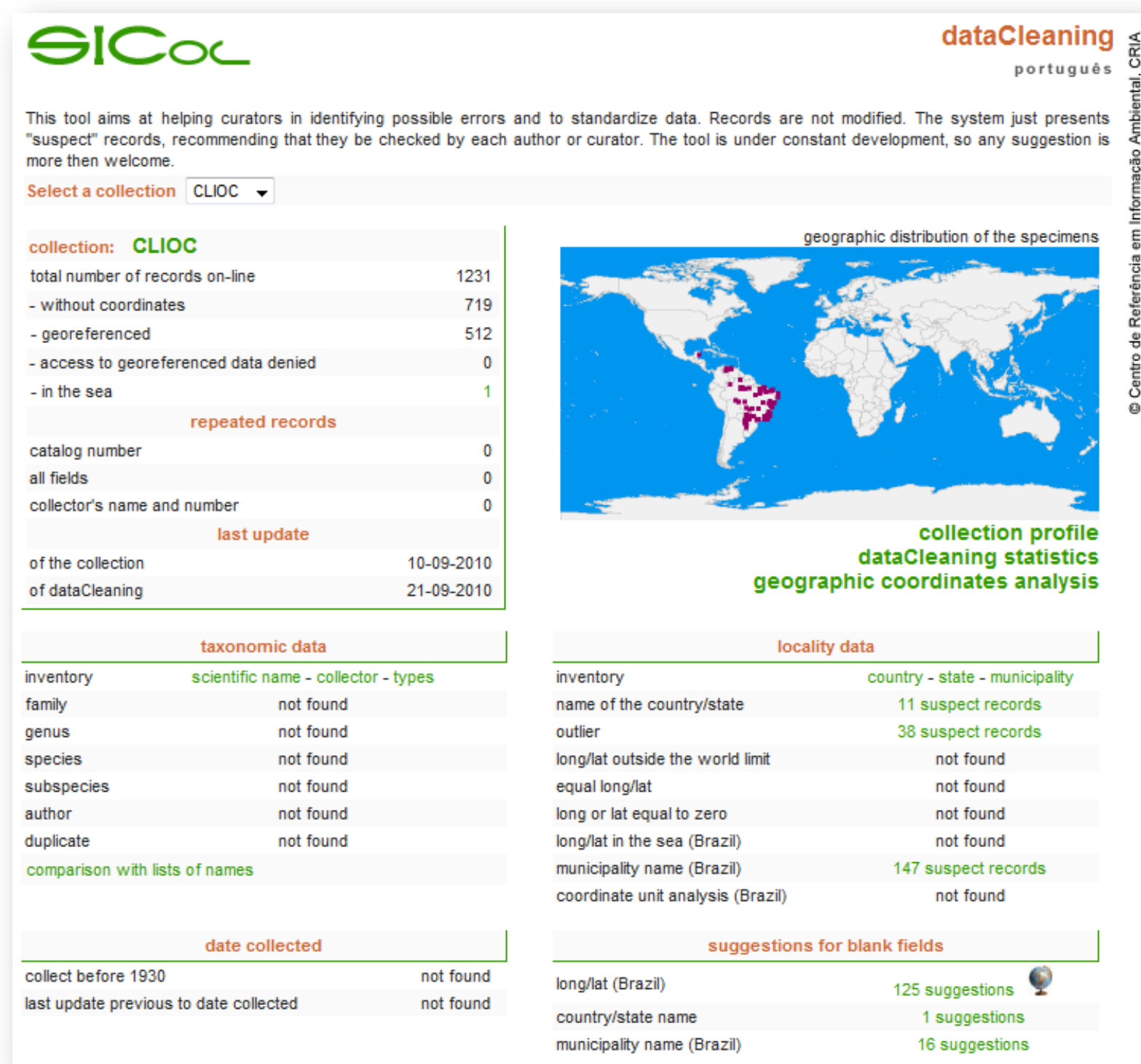**SIColNet: The Brazilian Microbial Resource Centre Network**

Challenges

SICol stands for Information System on Collections of Biotechnological Interest (Sistema de Informação de Coleções de Interesse Biotecnológico). The goal is to desseminate information on biological resources and specialized services provided by microbial collections in Brazil.
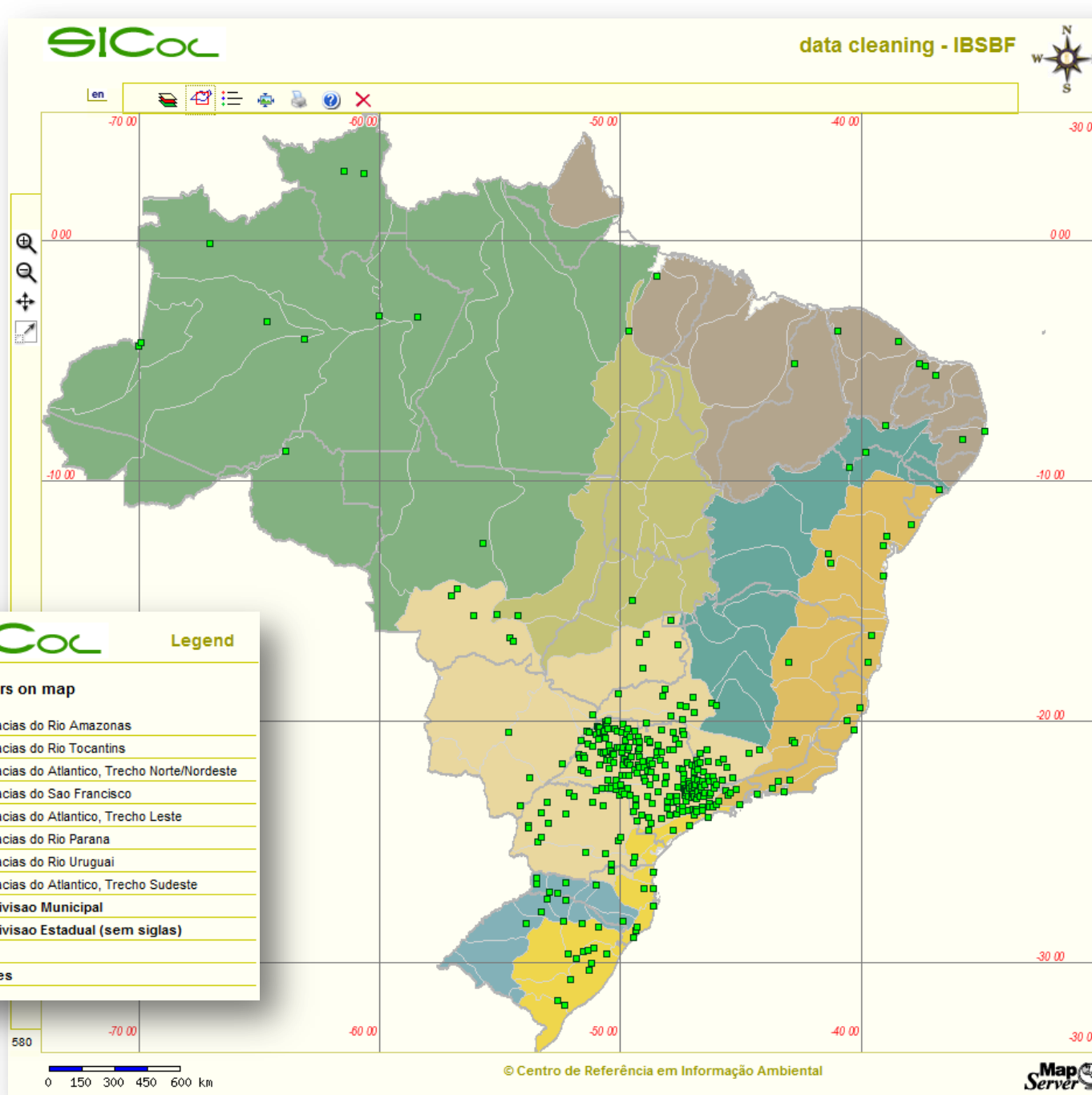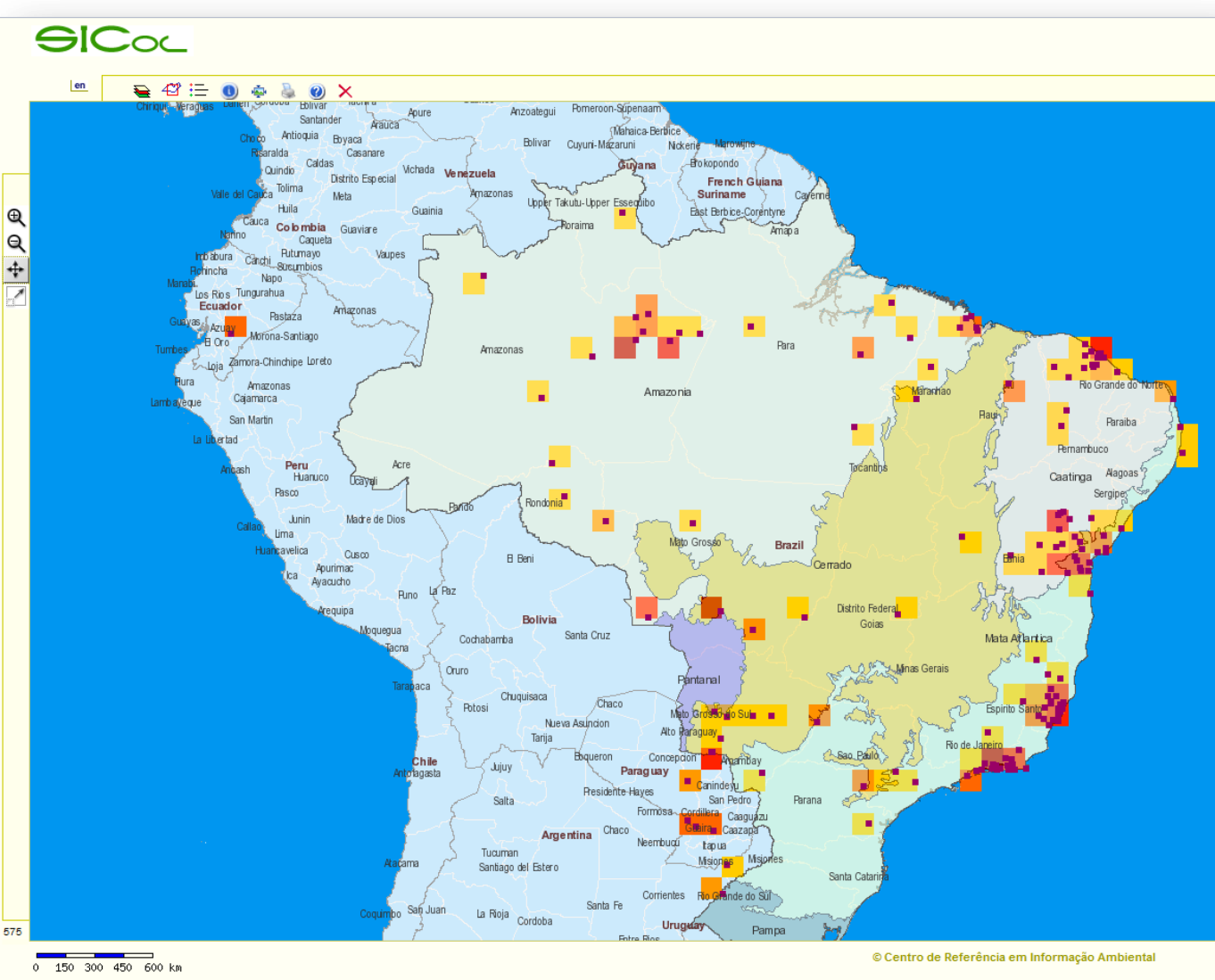
Architecture

Data cleaning

This tool "flags" potential errors, indicating "suspect records" to be checked by data providers. Main features include checking names using phonetic comparisons and the Catalogue of Life; checking dates indicating earliest collection and latest update; and checking geographic inconsistencies, country/state/municipality versus geographic coordinates.
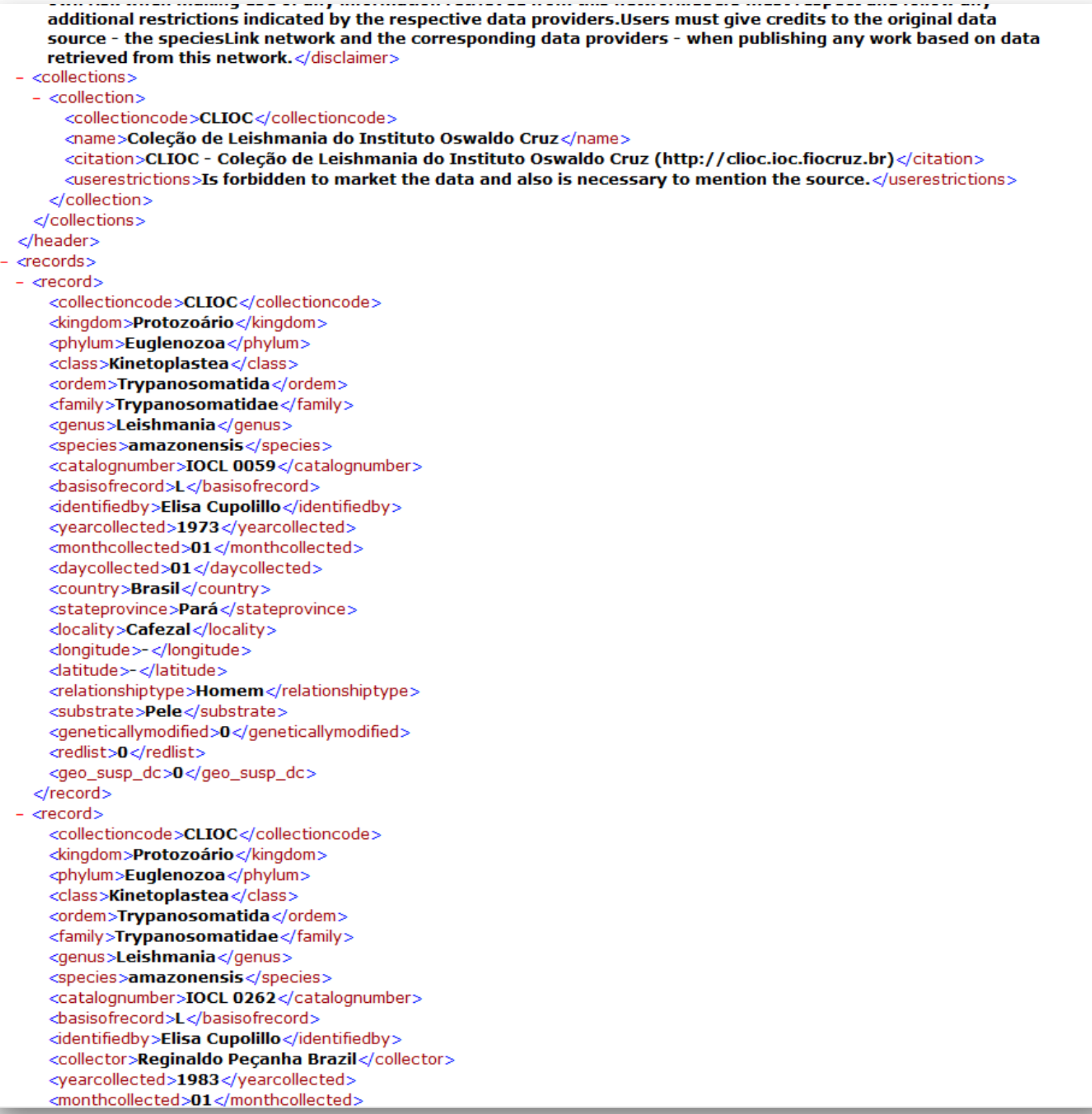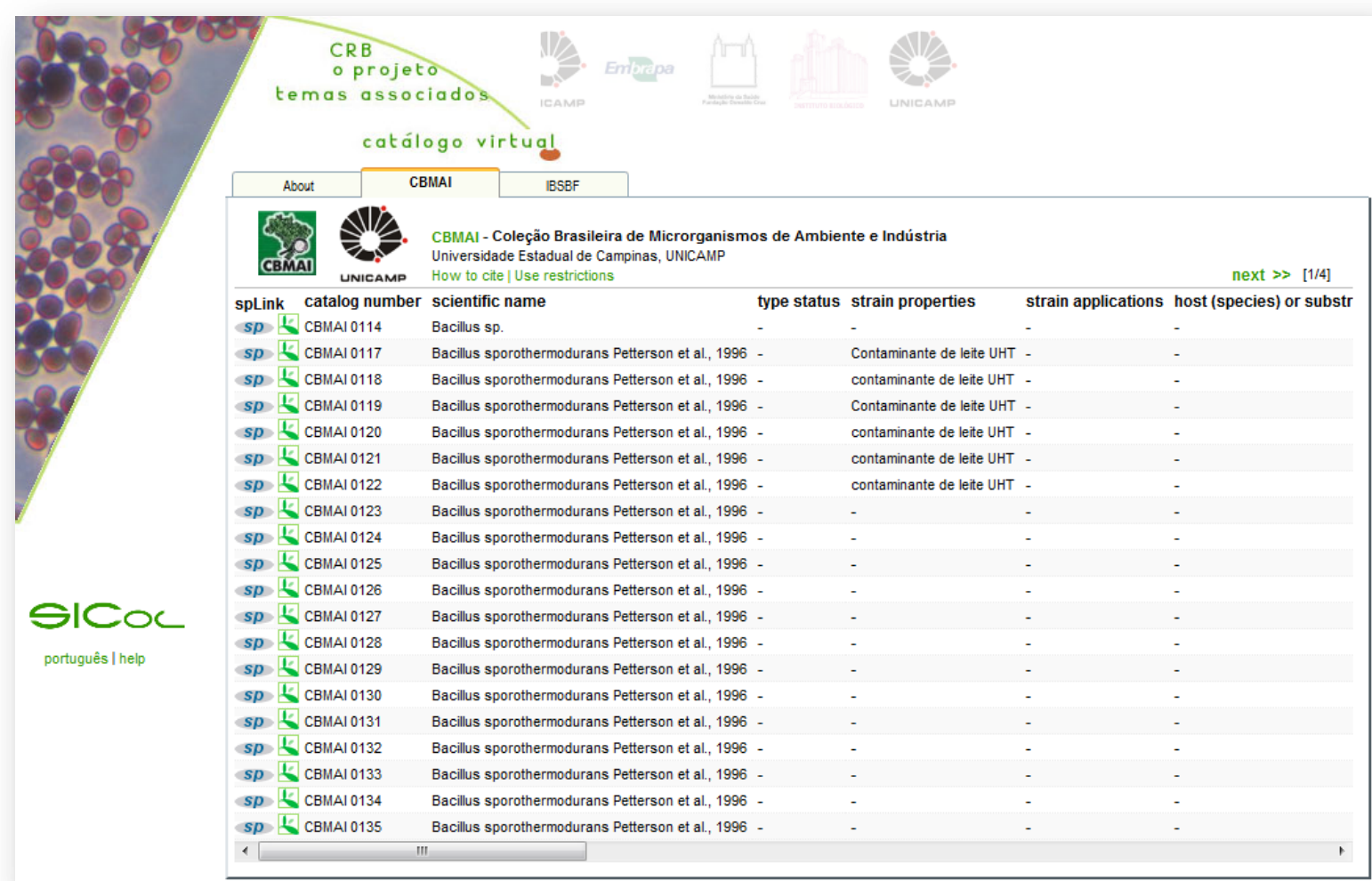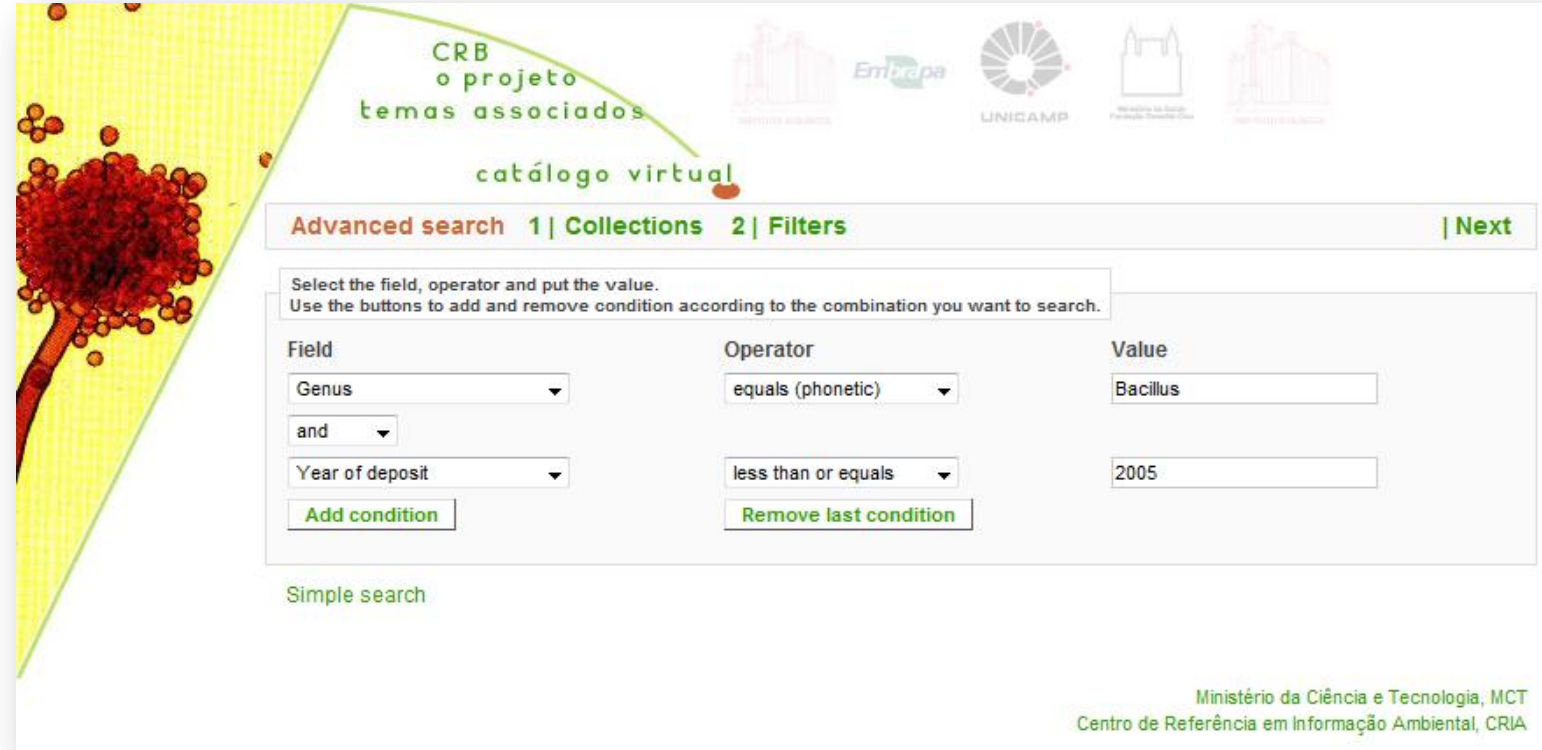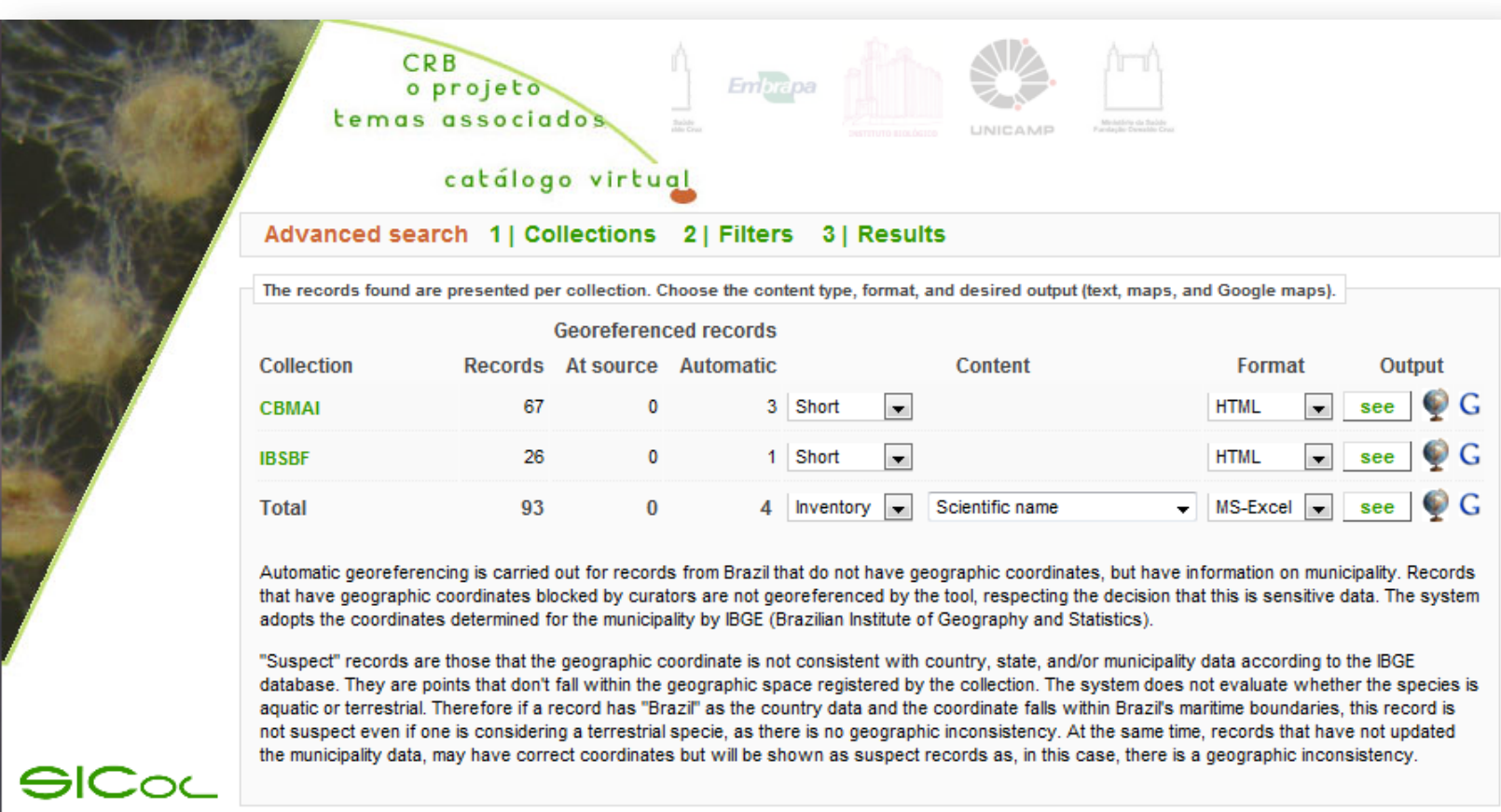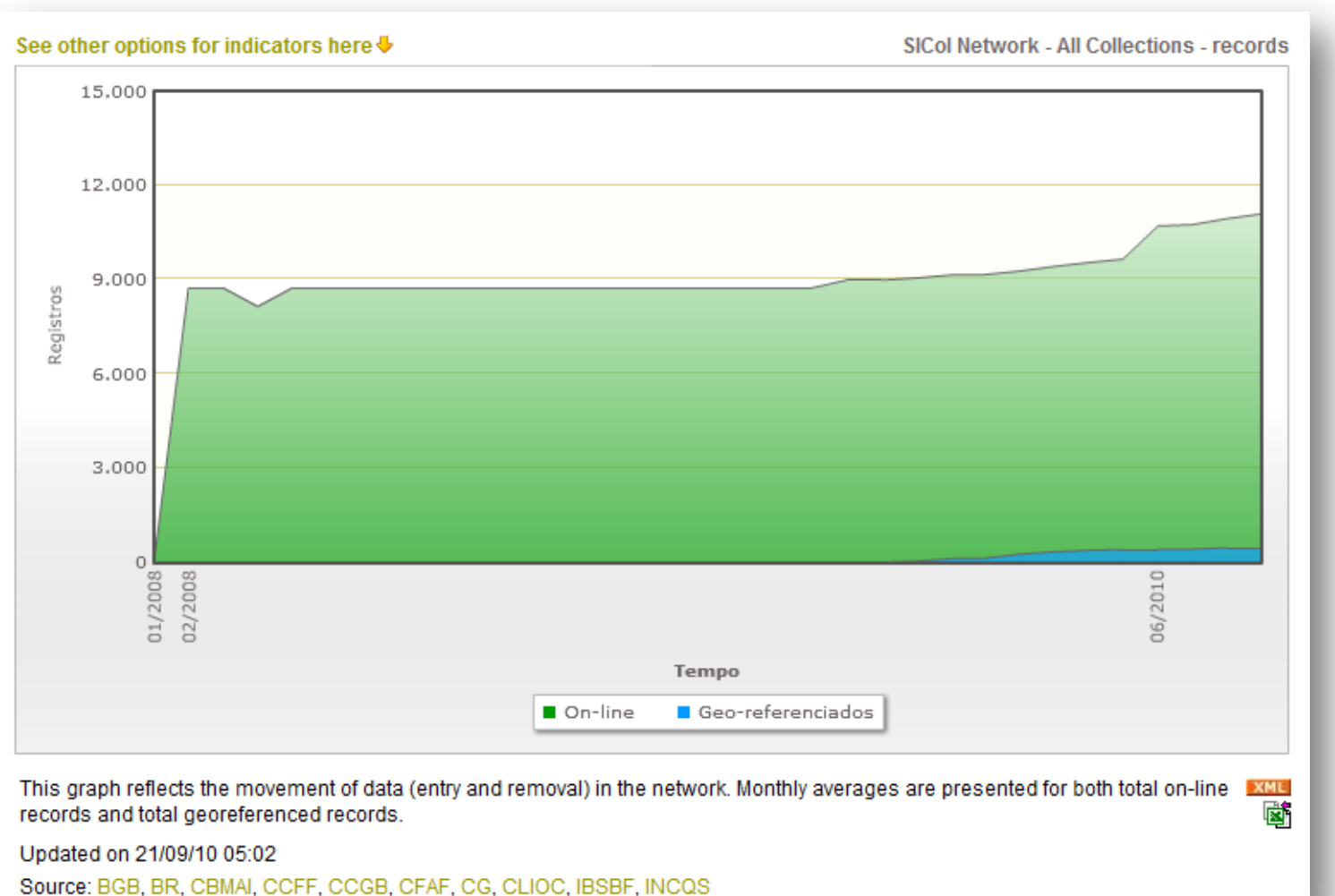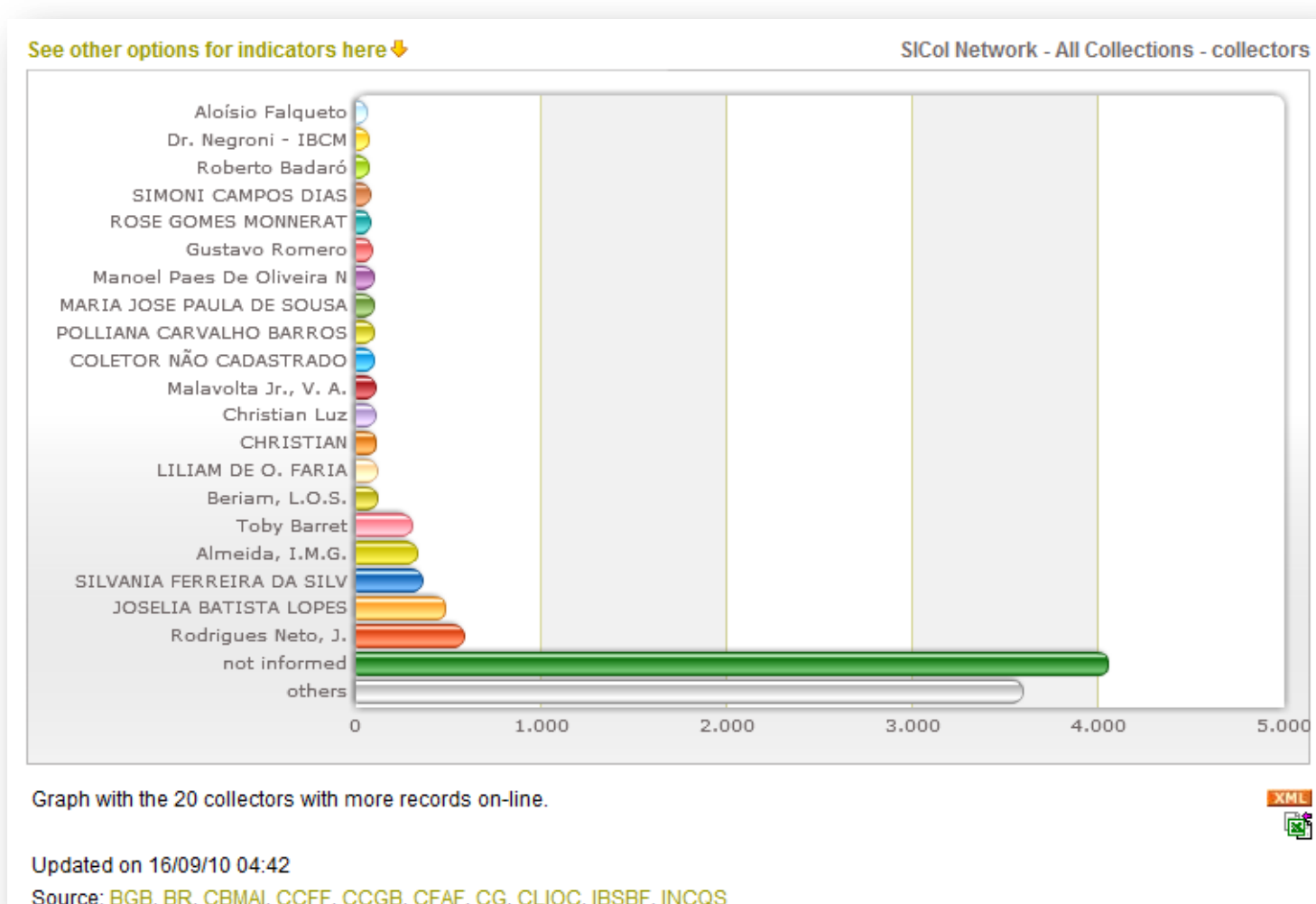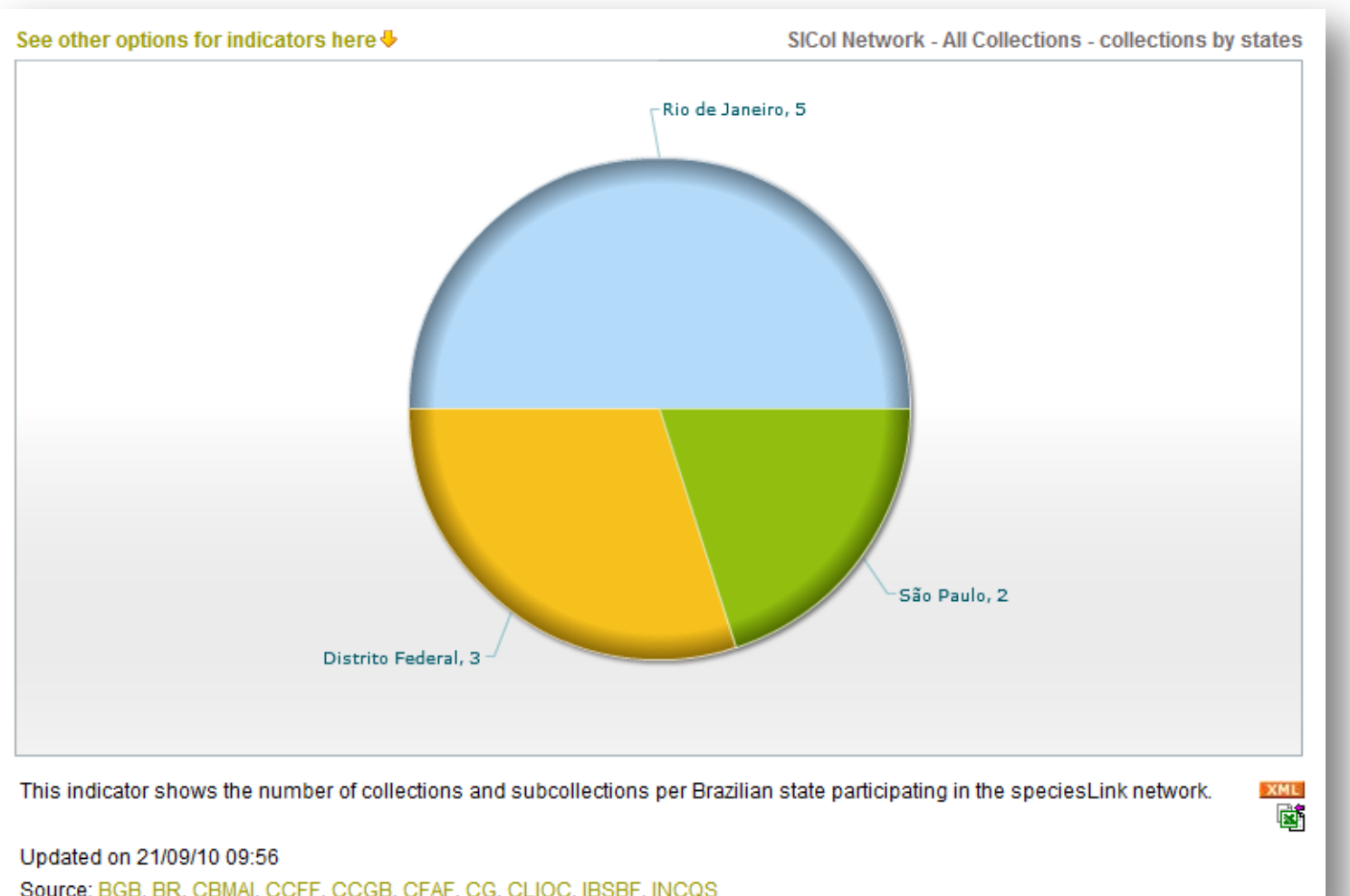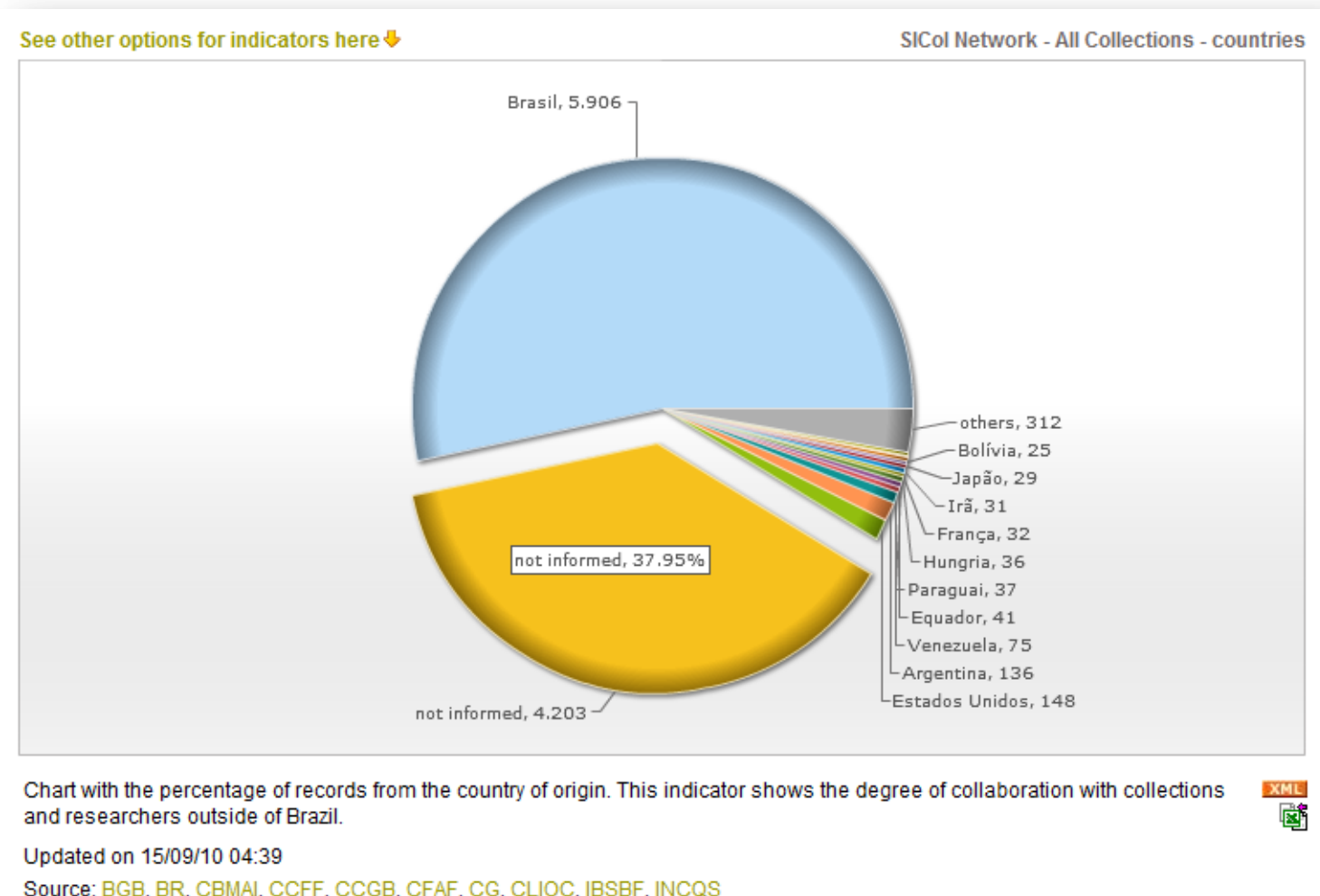

Centralized query

Users can select data source through a classification system. Data outputs include files (html, xml, excel), inventories, mapCRIA, Google maps. Data filters may be applied, such as type specimen, georeferencing at source or automatic. Results are presented per collection, integrated with red lists, geographic inconsistencies are highlighted, and smart links to CRIA's name server and GBIF are included.

Indicators

All indicators presented are based on on-line data. They are dynamic or daily reports presented as charts or diagrams. Indicators only reflect online data, so may not reflect the reality of each collection or group of collections, especially those with a low percentage of online data. As more collections are added to the network and more data is digitized, this tool may help indicate taxonomic and geographic information gaps.

Conclusions

The system enables data providers to have full control over their data, determining what can be openly shared. Data cleaning tools were designed to help collections improve data quality by indicating possible typing errors and data inconsistencies. Data can only be altered by each data provider at his/her own system. Data providers are fully acknowledged and have responsibility over the data that is shared.

http://sicol.cria.org.br
http://sicol.splink.org.br