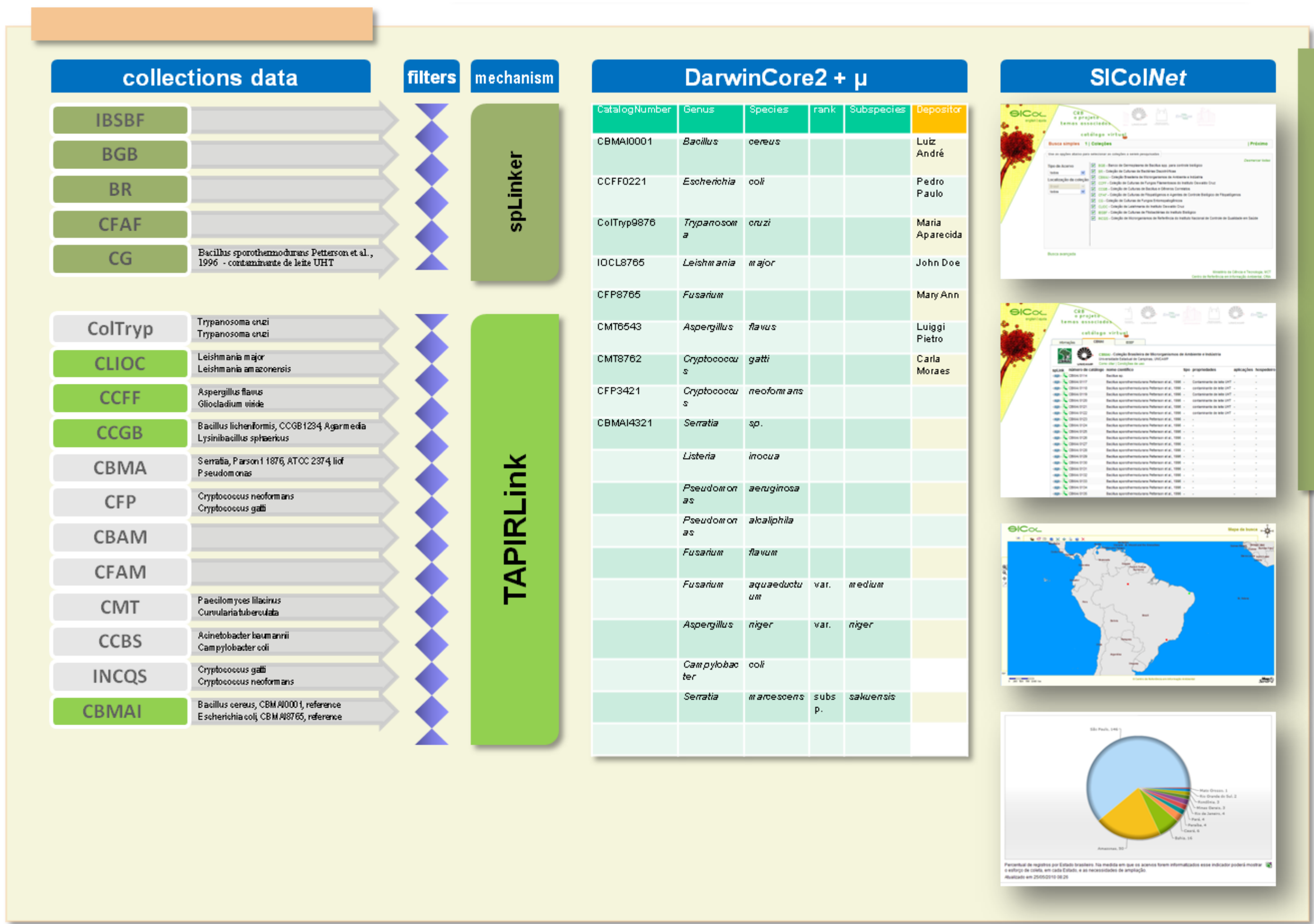


Alexandre Marino, Diego Rocha, Sidnei de Souza, Dora Ann Lange Canhos, Vanderlei Perez Canhos  
Centro de Referência em Informação Ambiental, CRIA, Brazil, [www.cria.org.br](http://www.cria.org.br)

## Challenges

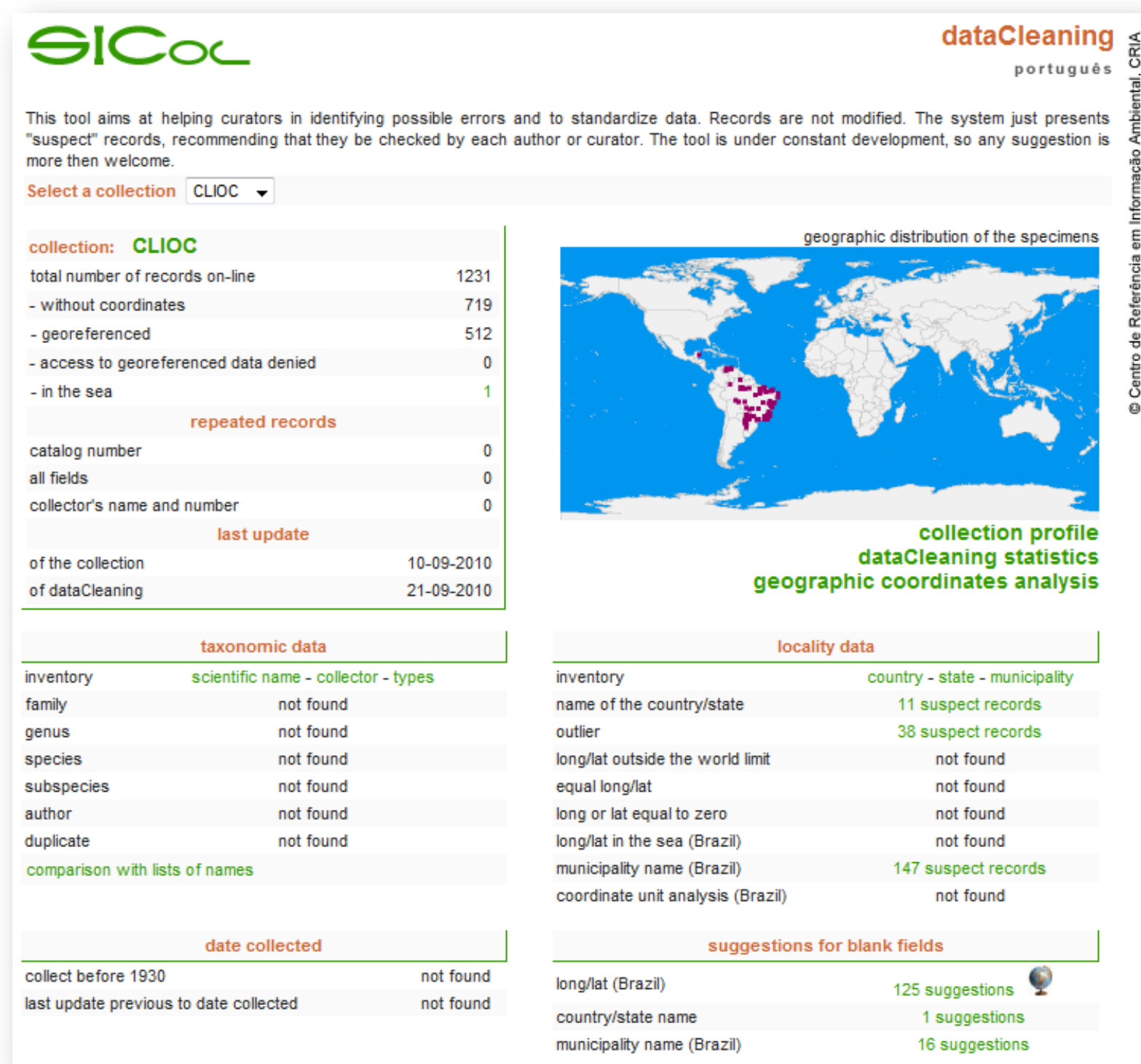
SICol stands for Information System on Collections of Biotechnological Interest (Sistema de Informação de Coleções de Interesse Biotecnológico). The goal is to disseminate information on biological resources and specialized services provided by microbial collections in Brazil.

## Architecture



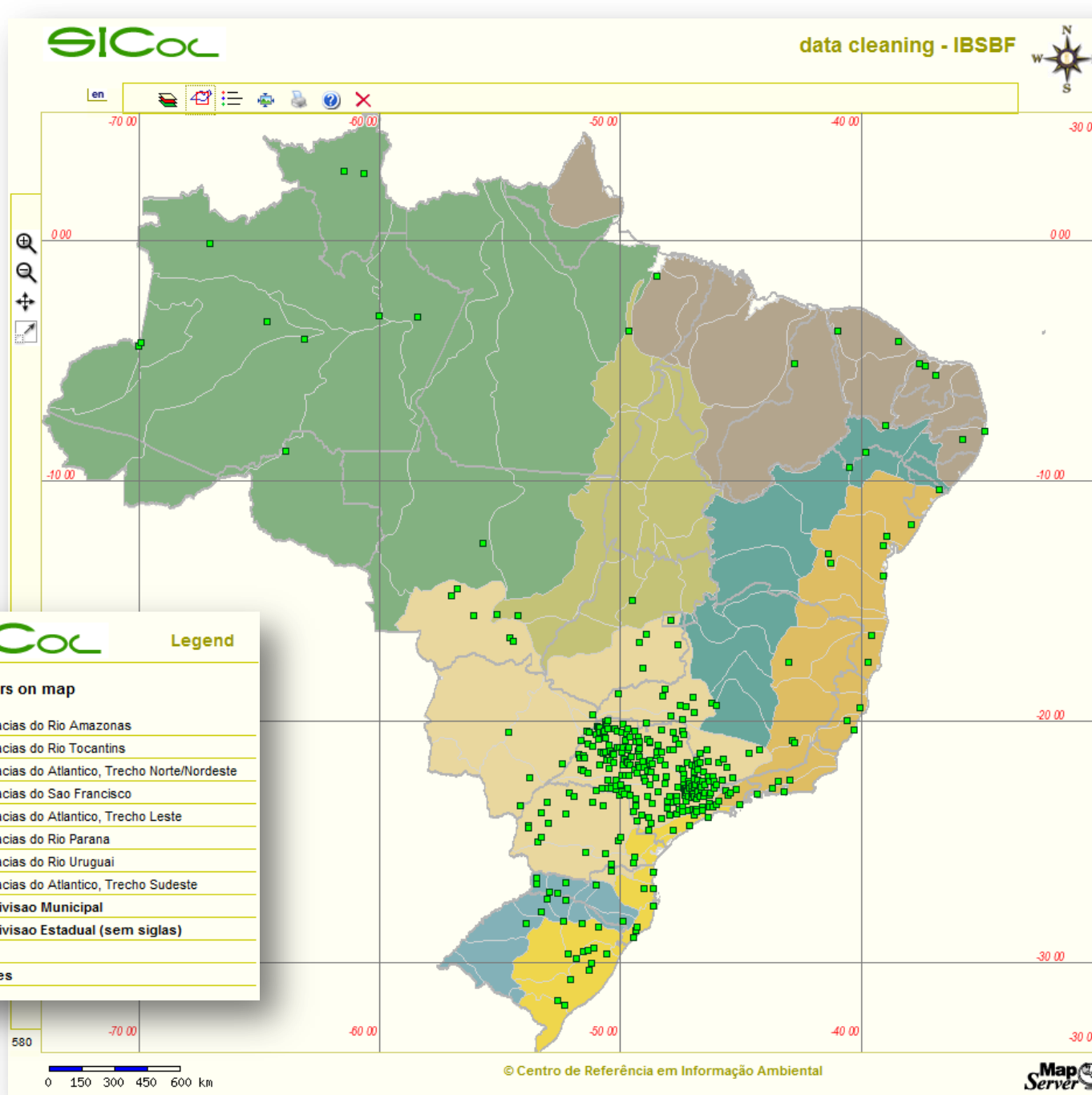
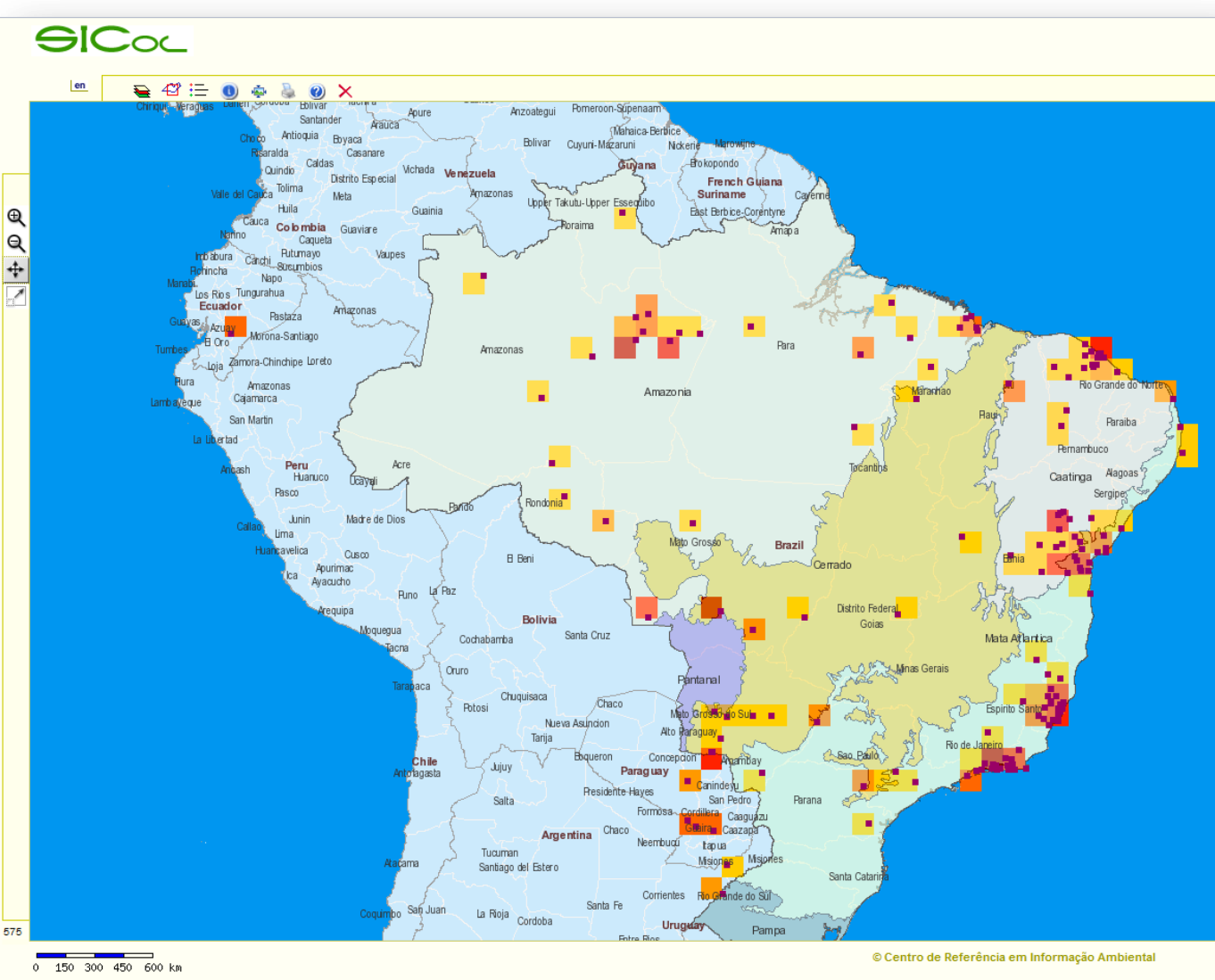
## Data cleaning

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.



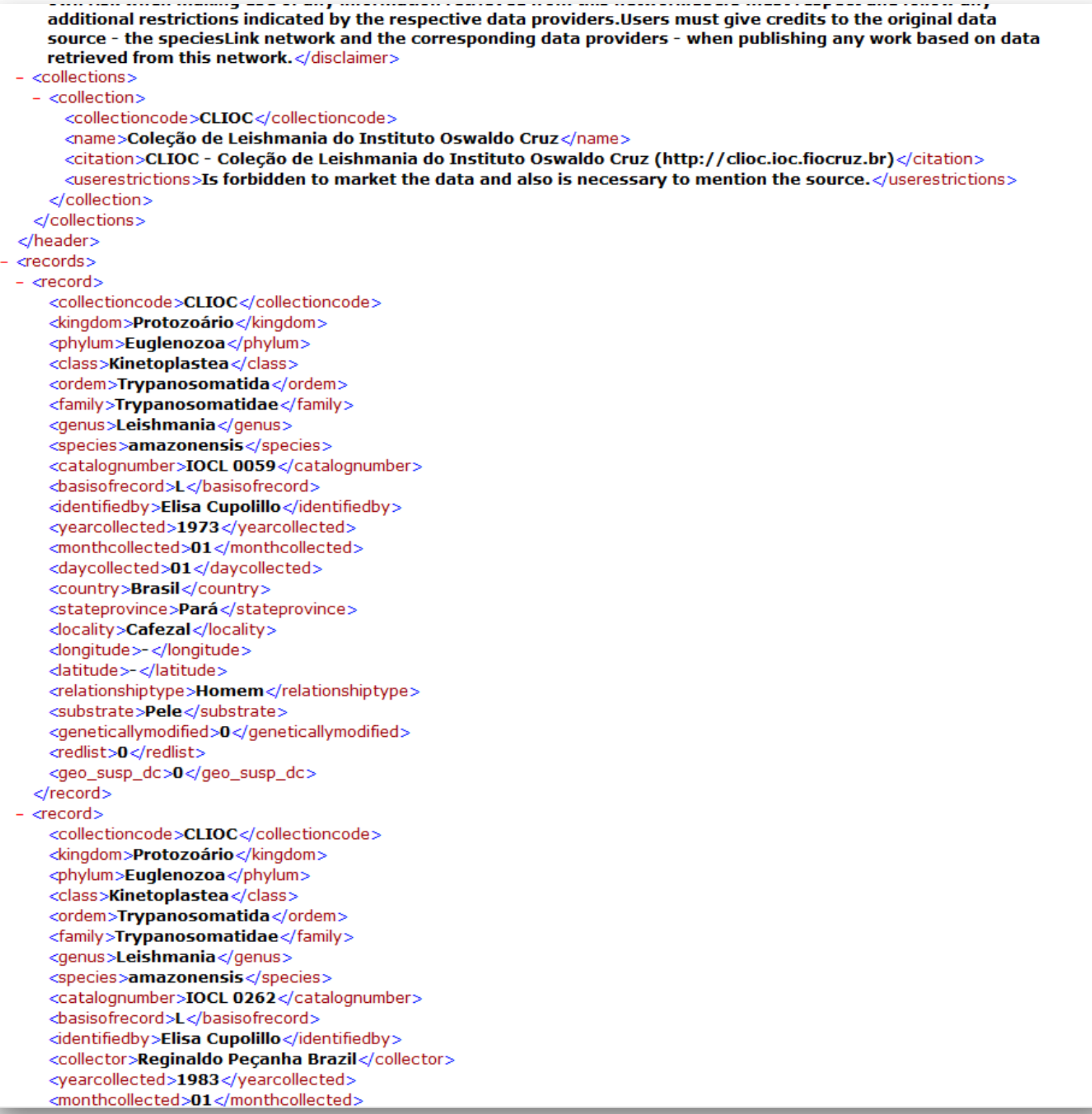
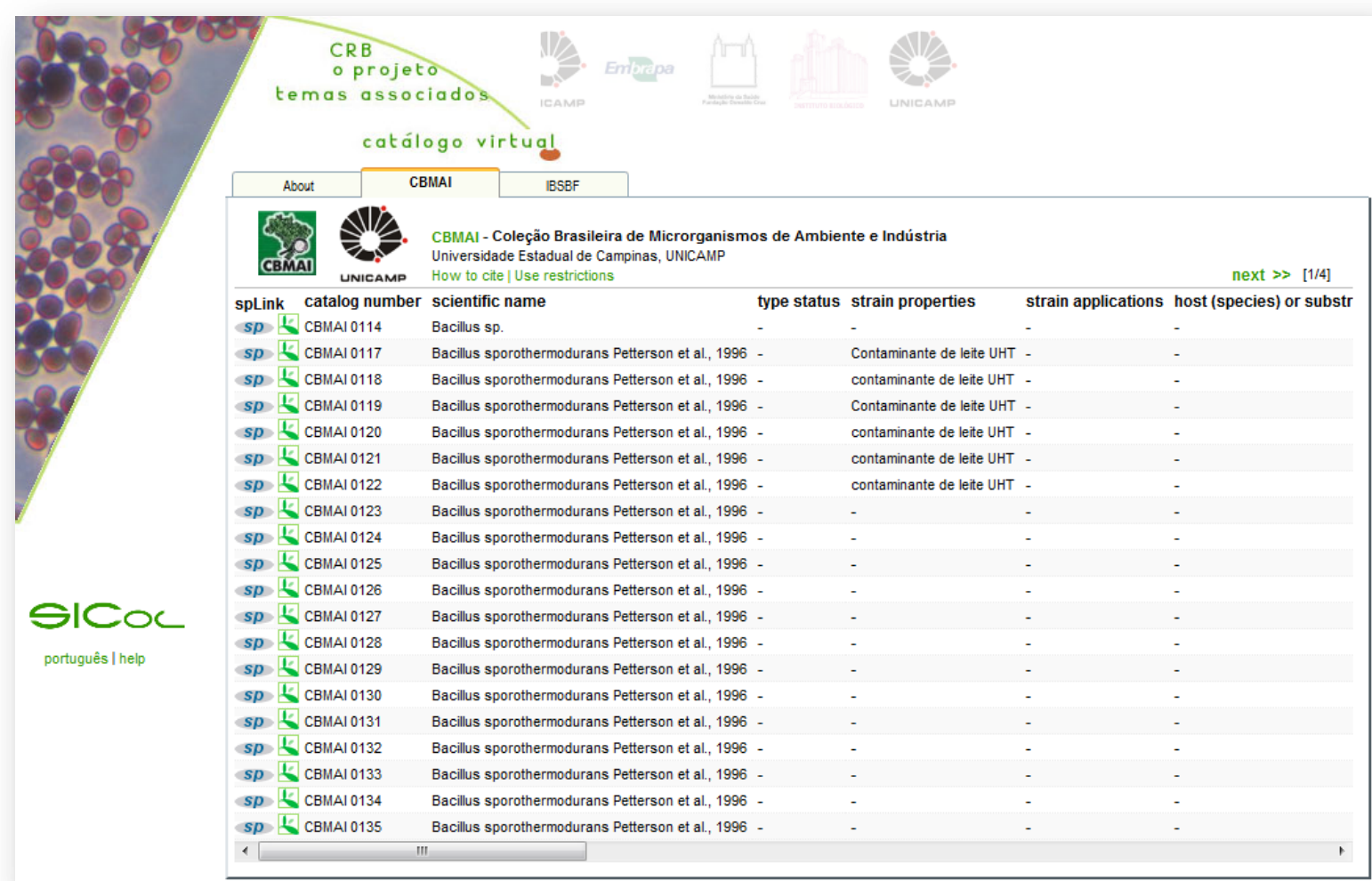
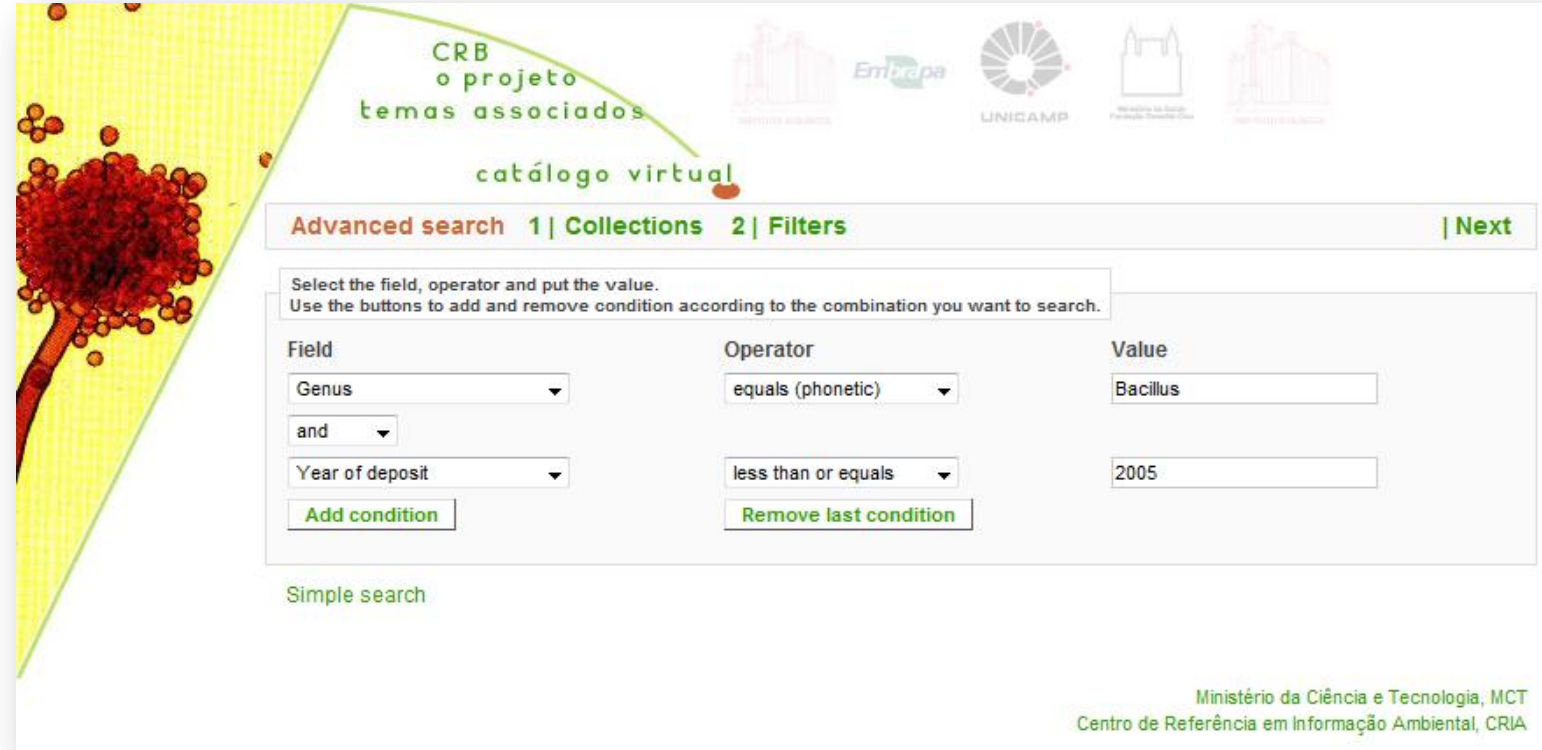
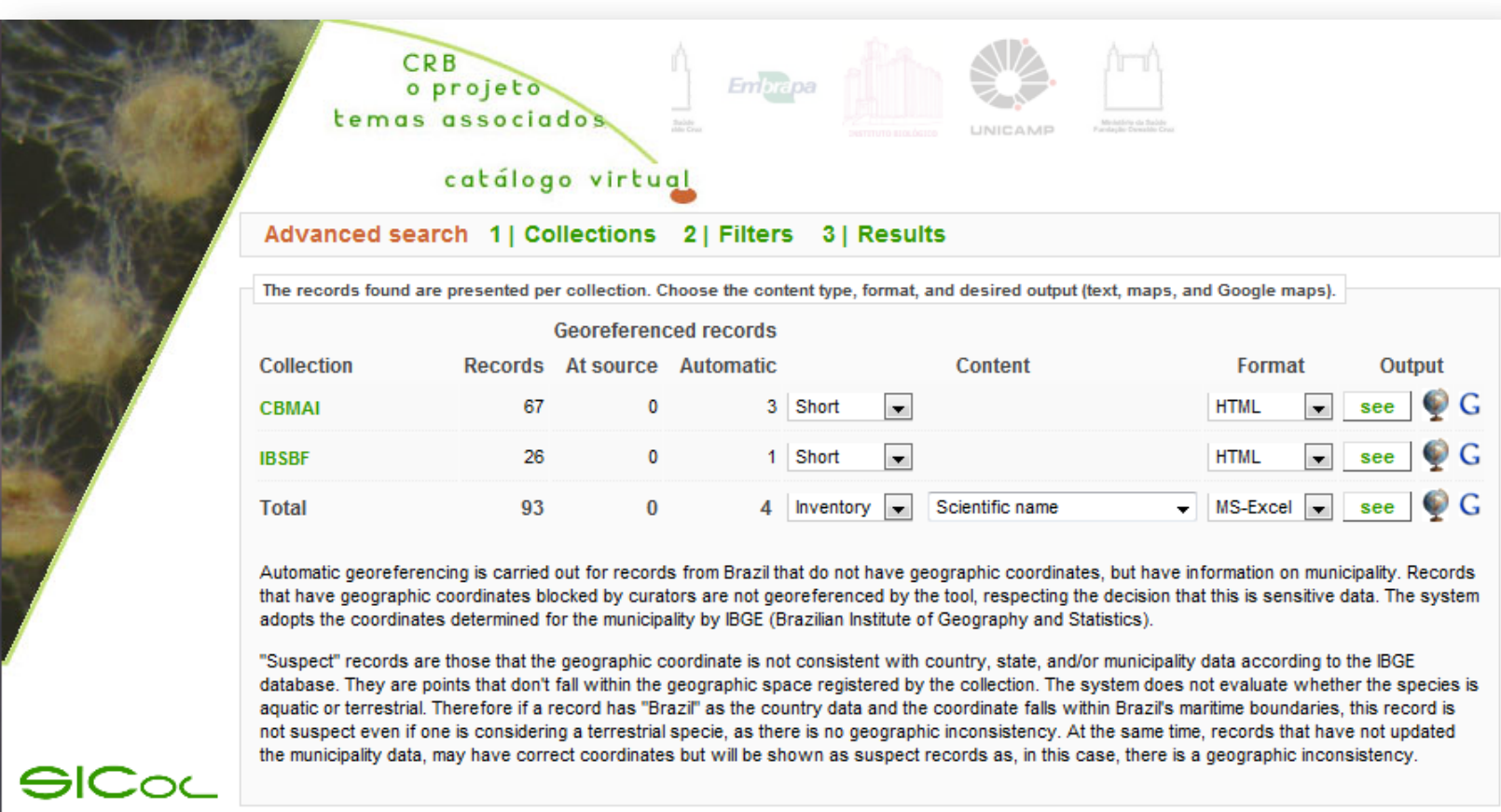
genus	species	subspecies	records	species link	status	catalog number
Trypanosoma	Trypanosoma	Trypanosoma	1	ver	1	618
Trypanosoma	Trypanosoma	Trypanosoma	3	ver	3	1599 1770 1961
Trypanosoma	Trypanosoma	Trypanosoma	1	ver	1	1444
Trypanosoma	Trypanosoma	Trypanosoma	1	ver	1	924
Trypanosoma	Trypanosoma	Trypanosoma	1	ver	1	1907
Trypanosoma	Trypanosoma	Trypanosoma	1	ver	1	1287
Trypanosoma	Trypanosoma	Trypanosoma	4	ver	4	1199 1515 1517 1520
Trypanosoma	Trypanosoma	Trypanosoma	335	ver	335	1744
Trypanosoma	Trypanosoma	Trypanosoma	1	ver	1	1
Trypanosoma	Trypanosoma	Trypanosoma	35	ver	35	2414 1709 2347 2415 2416 424 473 442 442
Trypanosoma	Trypanosoma	Trypanosoma	10	ver	10	1184 1709 2347 2415 2416 424 473 442 442
Trypanosoma	Trypanosoma	Trypanosoma	1	ver	1	185
Trypanosoma	Trypanosoma	Trypanosoma	1	ver	1	1312
Trypanosoma	Trypanosoma	Trypanosoma	2	ver	2	737 738
Trypanosoma	Trypanosoma	Trypanosoma	12	ver	12	1971
Trypanosoma	Trypanosoma	Trypanosoma	1	ver	1	1971
Trypanosoma	Trypanosoma	Trypanosoma	3	ver	3	2418 2419 2474
Trypanosoma	Trypanosoma	Trypanosoma	5	ver	5	2493 2494 2495 2496 2497

family	genus	species	subspecies	records
Trypanosomatidae	Trypanosoma	Trypanosoma	Trypanosoma	88
Trypanosomatidae	Trypanosoma	Trypanosoma	Trypanosoma	2
Trypanosomatidae	Trypanosoma	Trypanosoma	Trypanosoma	323
Trypanosomatidae	Trypanosoma	Trypanosoma	Trypanosoma	263
Trypanosomatidae	Trypanosoma	Trypanosoma	Trypanosoma	6
Trypanosomatidae	Trypanosoma	Trypanosoma	Trypanosoma	1
Trypanosomatidae	Trypanosoma	Trypanosoma	Trypanosoma	1
Trypanosomatidae	Trypanosoma	Trypanosoma	Trypanosoma	234
Trypanosomatidae	Trypanosoma	Trypanosoma	Trypanosoma	2
Trypanosomatidae	Trypanosoma	Trypanosoma	Trypanosoma	3
Trypanosomatidae	Trypanosoma	Trypanosoma	Trypanosoma	2
Trypanosomatidae	Trypanosoma	Trypanosoma	Trypanosoma	5
Trypanosomatidae	Trypanosoma	Trypanosoma	Trypanosoma	1
Trypanosomatidae	Trypanosoma	Trypanosoma	Trypanosoma	16
Trypanosomatidae	Trypanosoma	Trypanosoma	Trypanosoma	2
Trypanosomatidae	Trypanosoma	Trypanosoma	Trypanosoma	14
Trypanosomatidae	Trypanosoma	Trypanosoma	Trypanosoma	12
Trypanosomatidae	Trypanosoma	Trypanosoma	Trypanosoma	1
Trypanosomatidae	Trypanosoma	Trypanosoma	Trypanosoma	5
Trypanosomatidae	Trypanosoma	Trypanosoma	Trypanosoma	207
Trypanosomatidae	Trypanosoma	Trypanosoma	Trypanosoma	1
Trypanosomatidae	Trypanosoma	Trypanosoma	Trypanosoma	3
Trypanosomatidae	Trypanosoma	Trypanosoma	Trypanosoma	2
Trypanosomatidae	Trypanosoma	Trypanosoma	Trypanosoma	1
Trypanosomatidae	Trypanosoma	Trypanosoma	Trypanosoma	15



## Centralized query

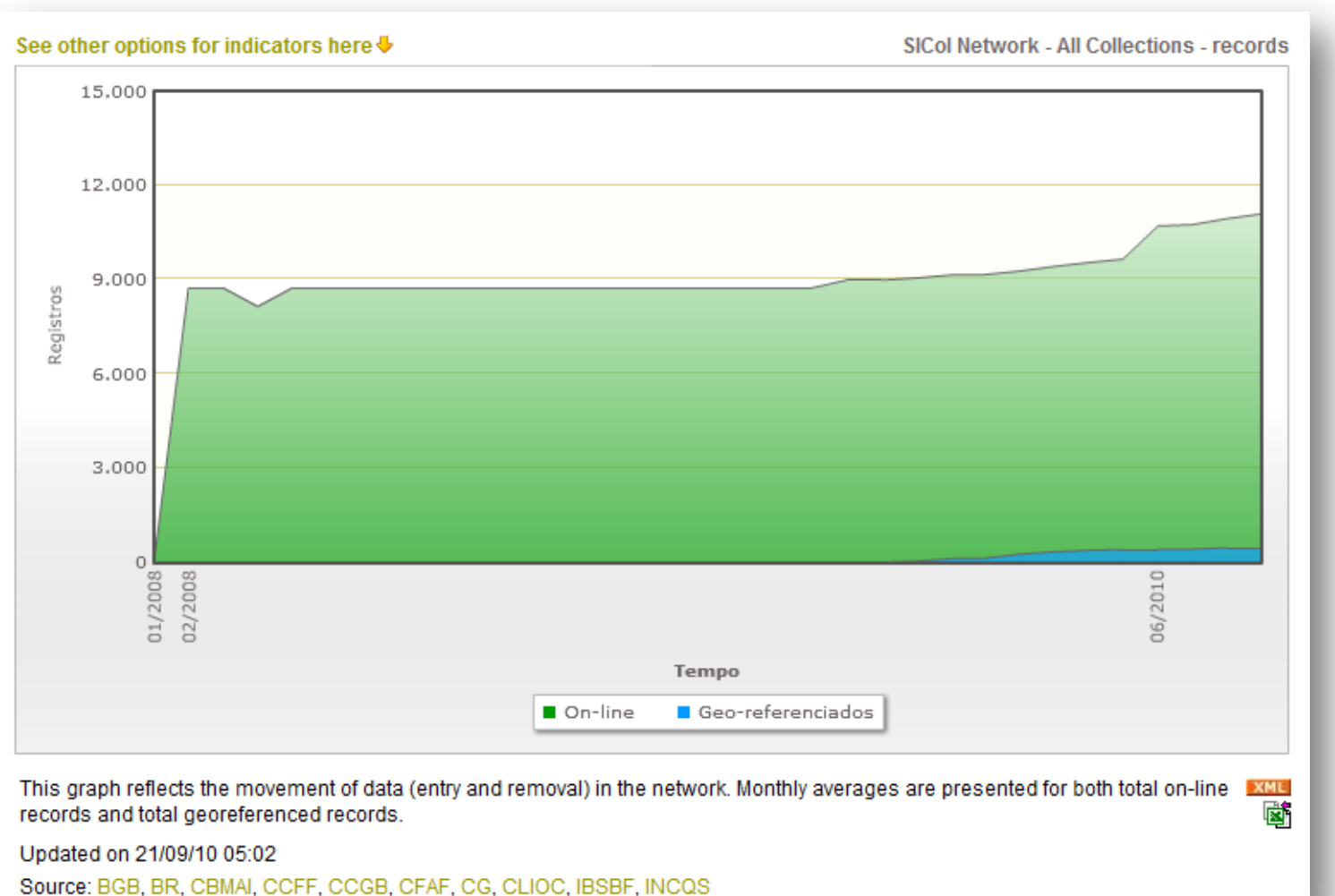
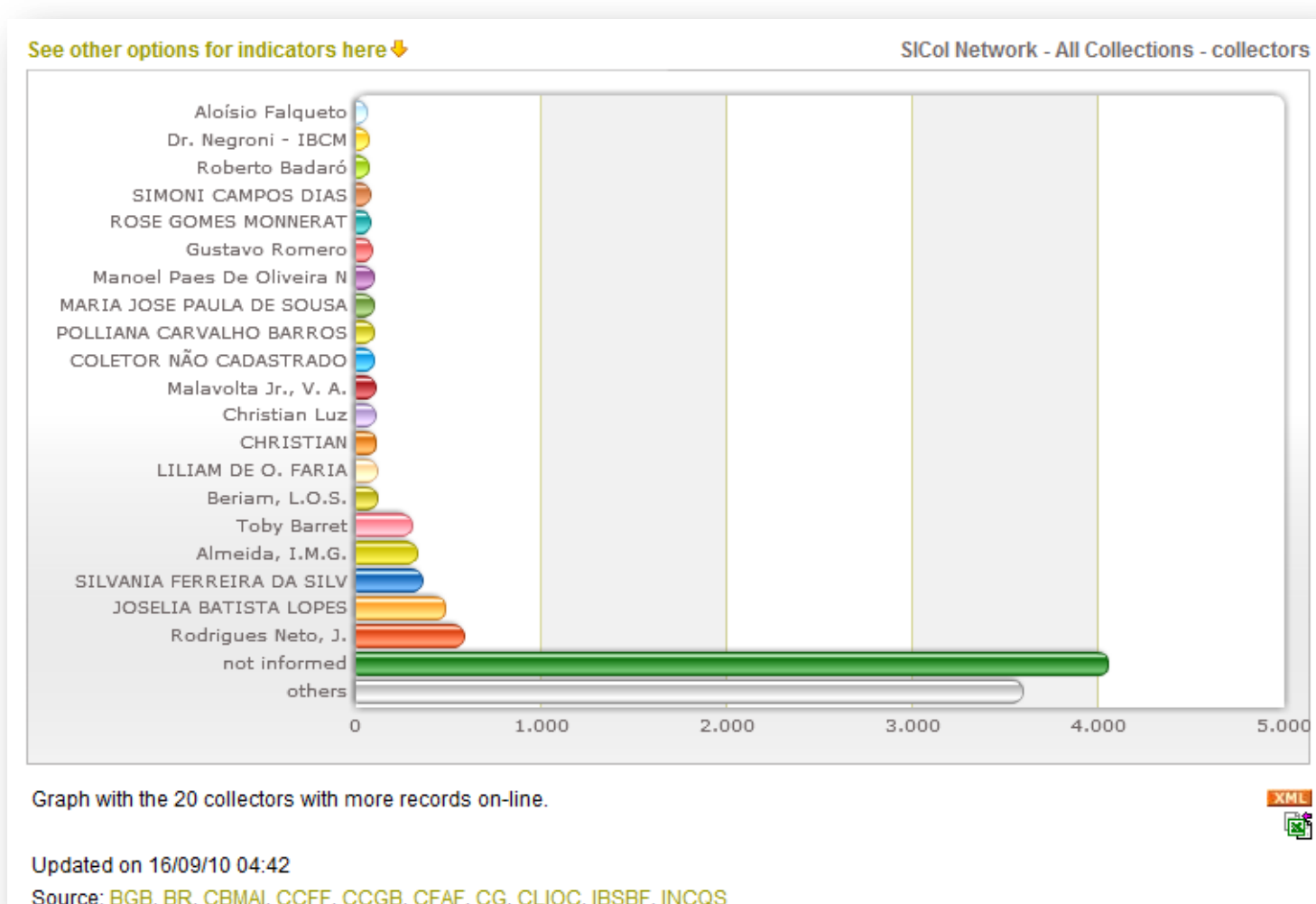
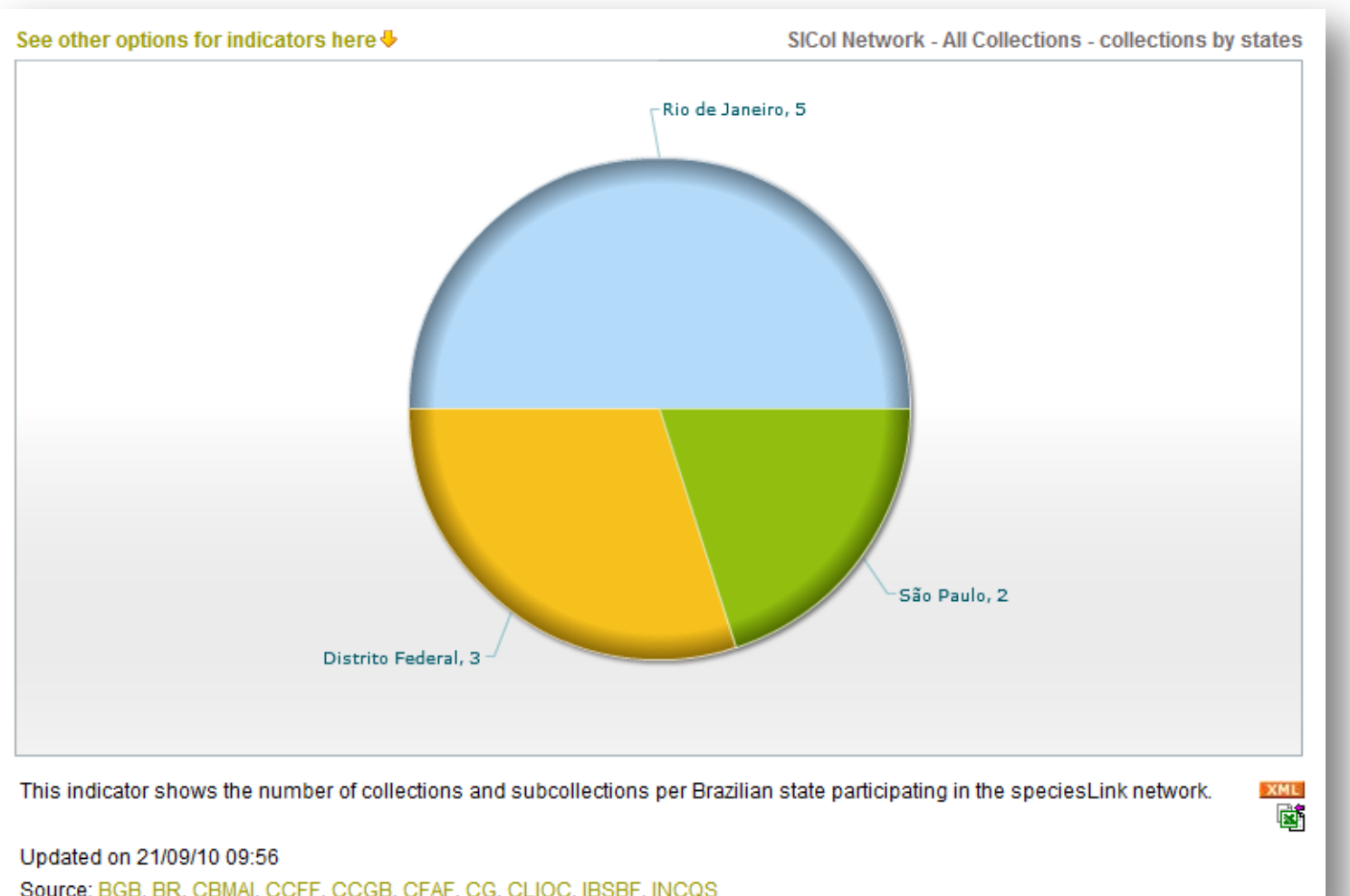
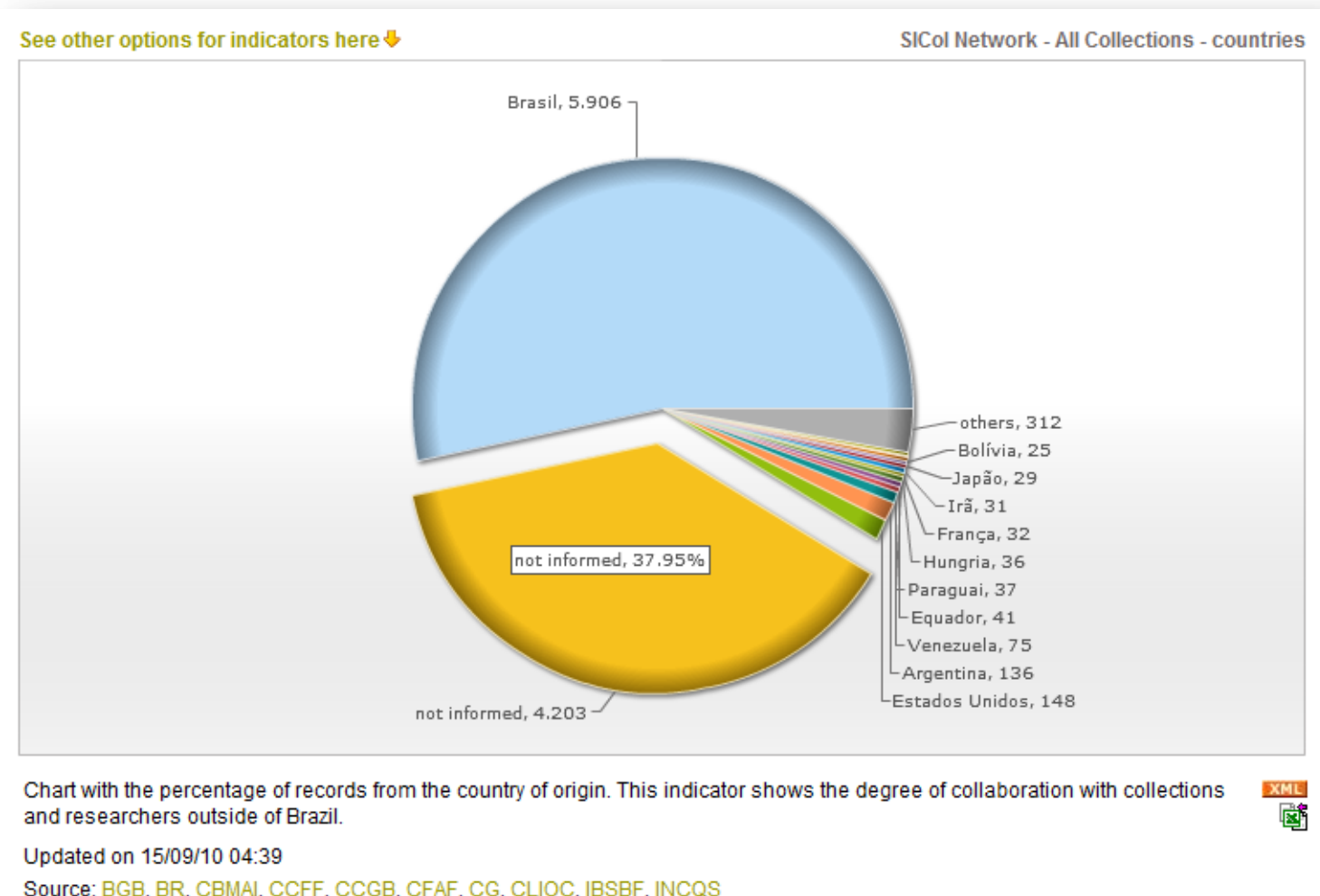
Users can select data source through a classification system. Data outputs include files (html, xml, excel), inventories, mapCRIA, Google maps. Data filters may be applied, such as type specimen, georeferencing at source or automatic. Results are presented per collection, integrated with red lists, geographic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



scientific name	total
Bacillus licheniformis	1
Bacillus sp.	22
Bacillus sporothermodurans Pettersen et al., 1996	41
Bacillus subtilis	17
Bacillus subtilis (Ehrenberg 1835) Cohn 1872	4
Bacillus thuringiensis	1
Bacillus thuringiensis subsp. darmstadtensis	1
Bacillus thuringiensis subsp. galleriae	1
Bacillus thuringiensis subsp. israelensis	1
Bacillus thuringiensis subsp. kurstaki	1
Bacillus thuringiensis subsp. tenebrionis	1
Bacillus thuringiensis var. thuringiensis	1
Bacillus (?)	1

## 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>