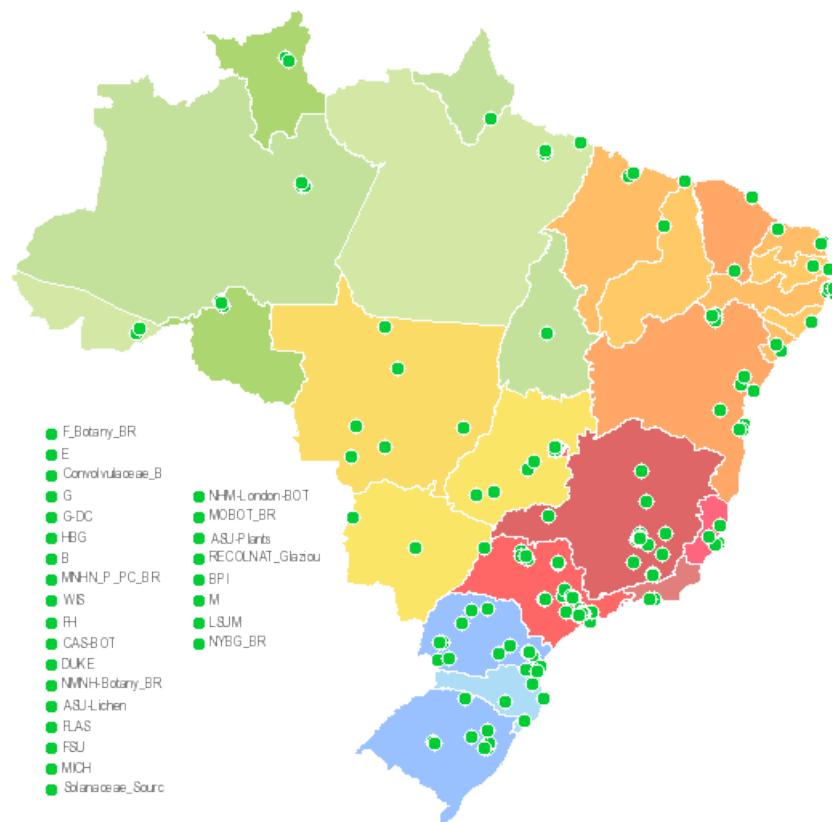# The impact of Brazil's Virtual Herbarium in e-Science

Reporting Period: February 2015 - February 2016

Suggested Citation: Canhos, Dora Ann Lange, Sidnei de Souza, Alexandre Marino and Vanderlei Perez Canhos. "The Impact of Brazil's Virtual Herbarium in E-Science." Midterm Progress Report for Open and Collaborative Science in Development Network (OCSDNet), February 2016.

Report Prepared by:

Dora Ann Lange Canhos

Sidnei de Souza

Alexandre Marino

Vanderlei Perez Canhos

BVH Steering Committee

Sumário

1. Executive Summary

The increase of knowledge on Brazilian biodiversity, associated with scientific advances to understand the evolutionary processes that generate and maintain this diversity, are fundamental to the sustainable use of this natural capital. Large investments have been made in developing cyber infrastructures to support research (Barjak et al., 2013). Examples from Brazil include Brazil's National Education and Research Network (RNP) and the National Centers for high performance processing (Cenapad). However, engineering breakthroughs alone will not be enough to achieve the outcomes envisaged for the undertaking of e-Science and other global collaborative activities supported by the cyber infrastructure. If it is to be achieved, it will more likely be the result of a nexus of interrelated social, legal and technical transformations (David, 2005; Tenopir et al., 2011).

The INCT – Virtual Herbarium of Flora and Fungi was established in 2009 to document, store, disseminate, and increase the knowledge base on the diversity of plants and fungi of Brazil. It adopted the *species*Link network as its cyber or e-infrastructure, focused on making data and tools openly and freely available online. Brazil's Virtual Herbarium (BVH) provides an ideal case study to help understand the drivers for collaboration and participation in this field.

The project's objective is to study both, the drivers for collaboration and the impact of the Brazilian Virtual Herbarium (BVH) in e-Science.

The project's focus during its first year was to evaluate whether data sharing has led to more recognition and support to data providers and what are the motivations to voluntary on-line collaboration.

Key observations for herbaria as data providers include:

- Data sharing increases visibility, which increases acknowledgement and support;
- Data sharing increases the interaction with students and visiting professors;
- Data sharing increases the number and variety of visitors to the herbaria;
- The network increases the importance and relevance of all participating herbaria;
- Internal organization was improved and overall planning and setting goals to be achieved was also enhanced as data was made available on-line; and,
- The human network established is the project's greatest asset

Although direct interaction between data providers and users through e-infrastructures is in its early development, for BVH it proved to be a powerful tool to improve data quality and to promote innovations, demanding new developments of the system. Improving data quality did not only include corrections concerning digitization errors but identification of specimens, proving that this is also a means of sharing knowledge.

As to crowd-sourcing through scientific workflows to produce and publish species' geographic distribution models, this has also proved to be a very important tool and strategy, resulting in public accessible models for 8% of all plants and fungi from Brazil.

This project carried out analysis that directly involve researchers that are contributing with data (herbaria) or with their knowledge (Annotation, BioGeo) which are normally overlooked by e-infrastructure that are focused on quantitative outputs, missing possible outcomes. This project brings another dimension to the work that is being carried out.

During 2016, new questions concerning usage will be investigated and the importance and reach of such an e-infrastructure shall be enhanced.

2. Research Problem

Research questions for the project's first year were:
- Has data sharing through the Brazilian Virtual Herbarium (BVH) lead to more recognition and support for data providers?
- What motivates crowd-sourcing within BioGeo and the Annotation System?

Although not included in the original proposal, BVH's Annotation system as a crowd-sourcing feature was included in the survey carried out in 2015.

3. Research Objectives and Emerging Findings

The research carried out throughout 2015, had two basic objectives:

a. Identify outcomes from on-line data sharing and,
b. Motivation and outcomes due to open and collaborative science through e-infrastructures

Identifying outcomes from on-line data sharing

One of the project's objectives is to identify possible drivers that motivate herbaria to openly share their data through an e-infrastructure and possible outcomes of this participation. A central research question of this project is "Has data sharing through the Brazilian Virtual Herbarium (BVH) led to more recognition and support for data providers?". The Brazilian Virtual Herbaria is one of the country's National Institutes of Science and Technology, a program of the Ministry of Science, Technology, and Innovation.

Methodology

Together with BVH's steering committee, the project team prepared a semi-structured questionnaire with both open and multiple-choice questions concerning possible outcomes from sharing data on-line.

A SWOT analysis was also carried out, requesting of each curator the Strengths, Weaknesses, Opportunities, and Threats concerning the Brazilian Virtual Herbarium.

The questionnaire was sent by email by the project's coordinator to curators of all 99 herbaria associated to the network and 39 answers were received. Each herbaria was classified into five separate groups, according to the size of its holdings:

- Up to 10 thousand vouchers;
- Between 10 and 50 thousand vouchers;
- Between 50 and 100 thousand vouchers;
- Between 100 and 200 vouchers; and,
- With more than 200 thousand vouchers.

The purpose of this "classification" was to evaluate if there were issues specifically related to the size of the herbarium.

All answers were tabulated and a report with the analysis of the answers was prepared and discussed with BVH's steering committee and presented at the evaluation and strategic planning meeting held in Belo Horizonte in April, 2015.

Thirty-five herbaria were present at the meeting, and a new round of discussions was carried out in smaller groups and presented in plenary. All this material is the basis of this report.

Therefore, this report is the result of the opinions of 17 herbaria that answered the questionnaire and participated in the meeting in Belo Horizonte, 22 herbaria that answered the questionnaire but were not present at the meeting, and 18 herbaria that did not answer the questionnaire but were at the meeting. Therefore, this study includes the opinion of curators from 57 herbaria, which at the time represented 58% of all associated herbaria of the network.

Results and Discussions

Outcomes

Table 1. Answers to the questionnaire from curators indicating outcomes derived from sharing their data through the e-infrastructure BVH

Size		< 10.000	10.001 - 50.000	50.001 - 100.000	100.001 - 200.00	>200.000	All
Number of Herbaria	No.	27	39	17	9	7	99
Number of Answers	No.	12	12	7	5	3	39
	%	44%	31%	41%	56%	43%	39%
Greater institutional recognition	No.	11	10	3	5	3	32
	%	92%	83%	43%	100%	100%	82%
Greater involvement with graduate courses	No.	9	8	5	2	2	26
	%	75%	67%	71%	40%	67%	67%
Increase in the Number of Visits	No.	10	12	6	2	3	33
	%	83%	100%	86%	40%	100%	85%
Increase of the holdings	No.	11	8	6	2	3	30
	%	92%	67%	86%	40%	100%	77%
Increase amount of grants	No.	6	6	3	2	3	20
	%	50%	50%	43%	40%	100%	51%

39% of all herbaria associated to the BVH answered the questionnaire. Outcomes derived from sharing data through a public e-infrastructure included (1) greater institutional recognition; (2) greater involvement with graduate courses, (3) increased number of visits to the herbaria; (4) increase of the holdings; and, (5) increase of grants.

As to being recognized or acknowledged by their own institution, the answers show that larger

institutions are fully recognized, which makes sense as the costs are much greater and an annual budget must be secured. This is not necessarily true for small herbaria, mostly in universities, that do not even have an official position for curators. The lack of recognition of the work or even of the existence of these herbaria by the host institution was always stated as a major problem of smaller collections. Therefore, the result of the survey indicating that 92% of herbaria with holdings of up to 10 thousand vouchers stated that sharing their data through the e-infrastructure gave them more visibility and institutional recognition is an important outcome of the project.

An important aspect of the network is that 95% of the participating herbaria are associated to graduate courses. The use of data and tools available in the Virtual Herbaria have become a routine in graduate courses such as botany, taxonomy, and ecology. By organizing and publicizing data of its holdings, herbaria have become more involved with graduate programs. Once again, looking at the smaller herbaria with up to 10 thousand vouchers, one can see that their involvement with the graduate courses increased. Many also indicated that by exposing the data of small, but geographically specific holdings, they attracted the interest of students and specialists. With this, the number of visitors increased as did the number of new samples deposited in their herbaria. These are important outcomes directly influenced by sharing data through the e-infrastructure.

Another major problem for smaller herbaria is external funding. With greater visibility and, in many cases, by submitting proposals as a network, 50% of the smaller herbaria with holdings under 50 thousand vouchers were successful in receiving external grants. However, not only did the small herbaria benefit from sharing the data of their holdings in an open platform, larger herbaria also acknowledged a great impact in the number of visits, holdings, and grants. Larger herbaria also manifested that their internal organization was improved and overall planning and setting goals to be achieved was also enhanced as data was made available on-line. By sharing their data on-line and by using all tools available for analysis, herbaria could work on data quality and plan future collecting efforts.

The increase of the holdings, attested by 77% of the herbaria, can be attributed to its greater visibility, its increased involvement with graduate students, and the increased interest of specialists in visiting the area where the herbaria is based. Some herbaria answered that besides the increase of the number of visitors, these are more diverse – both from different fields of knowledge and from different geographic areas.

SWOT Analysis

Curators were requested to indicate what they considered were strengths, weaknesses, opportunities and threats concerning the Brazilian Virtual Herbarium. Strengths and weaknesses referred to actions within the control of the network and opportunities and threats referred to external factors that are not within the control of the network but are important elements that must be monitored by the project.

STRENGTHS

All herbaria emphasized as strengths the social network, the value of data sharing, and the financial, technical and scientific support available through the project.

Social Network

The social network established and strengthened throughout the project, promoted increased interaction between curators and technicians from different institutions. Answers indicated that

there was a change in the mindset of the professionals involved that now feel more valued and responsible for the project's achievements. Increased self-esteem and a constant search for improvement was also noted. The increased geographic coverage of the network, with the participation of small herbaria, was emphasized, as many of these are regional collections, whose copies are underrepresented in other collections. Curators also indicated increased collaboration with students and researchers from other courses and institutions, and the visit of foreign researchers.

Data Sharing

Open sharing of textual data and images was viewed as a strength due to the greater visibility and acknowledgement of the role and importance of herbaria. Outcomes such as greater institutional recognition and deposits of new material (graduate students and researchers) were once again mentioned. On-line organization of data and the availability of tools to help find errors and inconsistencies were also mentioned and contributed to the improvement of the quality of the data that is being shared. An important observation mentioned was that data organization and on-line sharing also increases the knowledge curators have of their own holdings and enables better planning and the elaboration of strategies to increase and improve these holdings.

Project Support

The existence of the project with the support of the Brazilian government (CNPq) with funds for grants, materials, equipment, and for courses and visits of specialists was pointed out as being fundamental for the organization, digitization, and improvement of the holdings.

The fact that the project was developed from existing initiatives was considered a strength. These initiatives are the Brazilian Network of Herbaria of the Brazilian Botanical Society; the *species*Link network developed by CRIA; and the Brazilian National Research and Educational Network (RNP).

WEAKNESSES

The most important weakness cited by all curators refers to human resources. Not only are they insufficient, but specialists that are retiring are not being replaced. Even though the grants to hire students and technicians to work on the organization and digitization of the collections were mentioned as a strength of the project, here it was stated that these grants are transitory and for limited periods.

The same applies to infrastructure and the necessity of more training programs. The project promoted yearly meetings at the Congress of Botany and also held 2 general meetings (the first with representatives of 70 herbaria and the second with 35) to present and evaluate the results and to help plan the future. Many herbaria indicated that it would be important to hold more such meetings and this way guarantee a more participatory process.

OPPORTUNITIES

The following points were emphasized as opportunities to be monitored by the steering committee:

- The possible continuation of the federal government's program of National Institutes of Science and Technology is seen as an opportunity for continuity.
- Making data freely and openly available on-line is seen as an opportunity for new research insights and for the advancement of e-taxonomy, valuing the role of herbaria.
- The possibility of hiring professionals that were trained throughout the project is also seen

as an opportunity to ensure the transfer and multiplication of acquired knowledge.
- The advancement of information and communication technology is also seen as an opportunity to enhance the quality of the content shared on-line and to increase the interaction between herbaria (data providers) and data users.

The greatest threat mentioned was the discontinuity of the project. Within this train of thought, another point was the duplicity of similar projects, as opposed to collaborating and networking with existing initiatives.

The economic situation of the country and the lack of long-term public policies to support such e-infrastructures were considered threats.

Comments

This document synthesizes the opinion of curators from 57 herbaria associated to the Brazilian Virtual Herbarium. According to the vision of these professionals, this initiative brought great benefits to the herbaria and to the country and should continue.

The human network that was established is considered the project's most important asset. The *visiting specialists program*" that used on-line data to identify the herbaria to be visited and specialists required, the courses and technical visits and remote support given were actions that strengthened this human network with a sense of being part of the group.

The capillarity of the network, with the participation of at least one herbaria from every state of the country, including small and regional herbaria, and the involvement with graduate courses are an important contribution of the project.

Therefore, returning to our central research question "Has data sharing through the Brazilian Virtual Herbarium (BVH) led to more recognition and support for data providers?", the answer is yes.

Motivation and outcomes due to open and collaborative science through e-infrastructures

The project developed two important mechanisms to allow users to collaborate directly to the e-infrastructure:
- An annotation system; and,
- A workflow (BioGeo) to enable users to produce and publish species' ecological niche models.

Questions include:
- What motivates users to send their comments (*Annotation System*)?
- What motivates researchers to produce and publish their models through *BioGeo*?
- Are there mechanisms that could be used to increase participation?

Annotation System

Within the Brazilian Virtual Herbarium project, feedback mechanisms were developed to allow users to send their comments about specific data records. When clicking on the "new comment" icon, a window pops up for users to provide their input (figure 1).

Figure 1. Popup window to enable users to send comments to curators

The system presents the full data record and users must fill out the form with their name and email, select the subject – scientific name, identification, geography, and other – and write their comment. When clicking on **send** the comment is sent to the email indicated by the user for confirmation. Once confirmed, the curator receives the email that is achieved in a database. Figure 2 shows a record with an associated comment.

Figure 2. Record RON 8523 with an associated comment.

Figure 2 shows that the herbarium RON has voucher 8523 identified as *Blechnum* and the comment indicates that it is *Thelypteris interrupta* (Willd.) K. Iwats. (Thelypteridaceae). This comment was sent on October 10, 2015, but the last time this database was updated was in June 2015. Even though the record was not altered by the curator, it appears as an on-line annotation associated to the specific record. Users can therefore benefit from a specialist comment even before the data has been revised and altered. Users can also check this information as there is an image associated to the record.

Methodology

When this survey was prepared, the system had received 622 comments from 141 people. 473 comments referred to the scientific name, 68 to the identification of the material, 59 to the geographic data, and 22 classified as "other". 85.5% of the comments referred to data records of plants, 13.5% to animals and 1% to microorganisms.

An email was posted to all 141 users who sent their comments through the annotation system asking "What was the motivation for using this tool?". In order to facilitate the analysis, six options were offered:

(1) Contribute to the improvement of the quality of the data;
(2) Correct errors in order to enable the use of the data in their research;

(3) Correct errors in order to use the data in the BioGeo workflow;
(4) Check the determination and/or geographic information to use this information in the List of Species of the Brazilian Flora;
(5) Check the determination and/or geographic information to use this information in the red list assessment (CNCFlora);
(6) Others. In this option, users were asked to specify what other reasons they had.

People could choose more than one option. We also asked whether the herbarium accepted their comments and corrected possible errors, asking them to choose one of the four options below:
- All records were corrected
- Most records were corrected
- Some records were corrected
- No record was corrected

Results

Of the 141 emails sent, we received 20 answers, around 14% of the total.
- 85% indicated that their motivation was to contribute to the improvement of the quality of the data
- 50% to correct errors in order to enable the use of the data in their research
- 5% to correct errors in order to use the data in the BioGeo workflow
- 5% to check the determination and/or geographic information to evaluate the species' distribution and include this information in the List of Species of the Brazilian Flora

No one indicated the use of the tool to use the data in the red list assessment and no other motivation was included.

As to whether, to their knowledge, the collections are benefiting from their comments to correct possible errors, only 16 of the 20 specialists answered this question.
- 15% indicated that all records were corrected
- 15% indicated that most records were corrected
- 15% indicated that some records were corrected
- 15% indicated that the records were not corrected
- 20% indicated that they do not know whether the data was corrected

Comments

The most important driver for participation is to contribute to improve data quality and to enable the use of the data in research. It may be that the 20% that did not answer the second block did not check to see whether the records were modified. However, we can conclude that 60% did not only contribute with their comments but also checked to see if the data was modified.

BioGeo (Biogeography of the Flora and Fungi of Brazil)[1]

BioGeo is a system developed to expand the knowledge about biogeography of plants and fungi of Brazil, using modeling techniques of potential distribution and counting with an active participation of specialists. A diagram of the workflow is presented in figure 3.
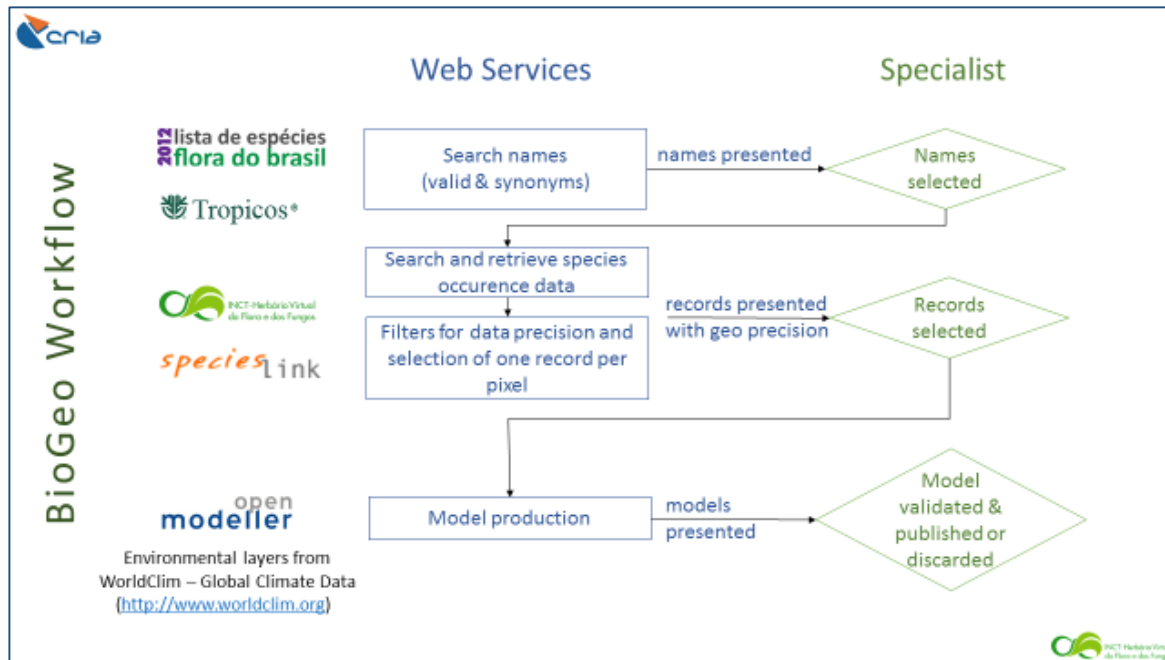

Figure 3. BioGeo Workflow

When a specialist registers in the system, he/she indicates the species or genera that he/she wants to model. The process begins when a specific species is selected. Through the workflow, the system using web services for the List of Species of the Brazilian Flora (provided by CRIA) and Tropicos (provided by Missouri Botanical Garden) presents a list of names (valid names and synonyms) to the specialist, who selects those to be included for searching. The names are searched through speciesLink's web services and the results go through a filter for data precision (lat/long) that selects one record per pixel. Records selected by the system and all other records found are presented to the specialist who defines which records will be used in modeling. Depending on the number of point data, different algorithms are used and models are produced using the openModeller web services and WorldClim data. The resulting models, together with a consensus model, are presented to the specialist who then decides whether it should be published or discarded (Figure 4).
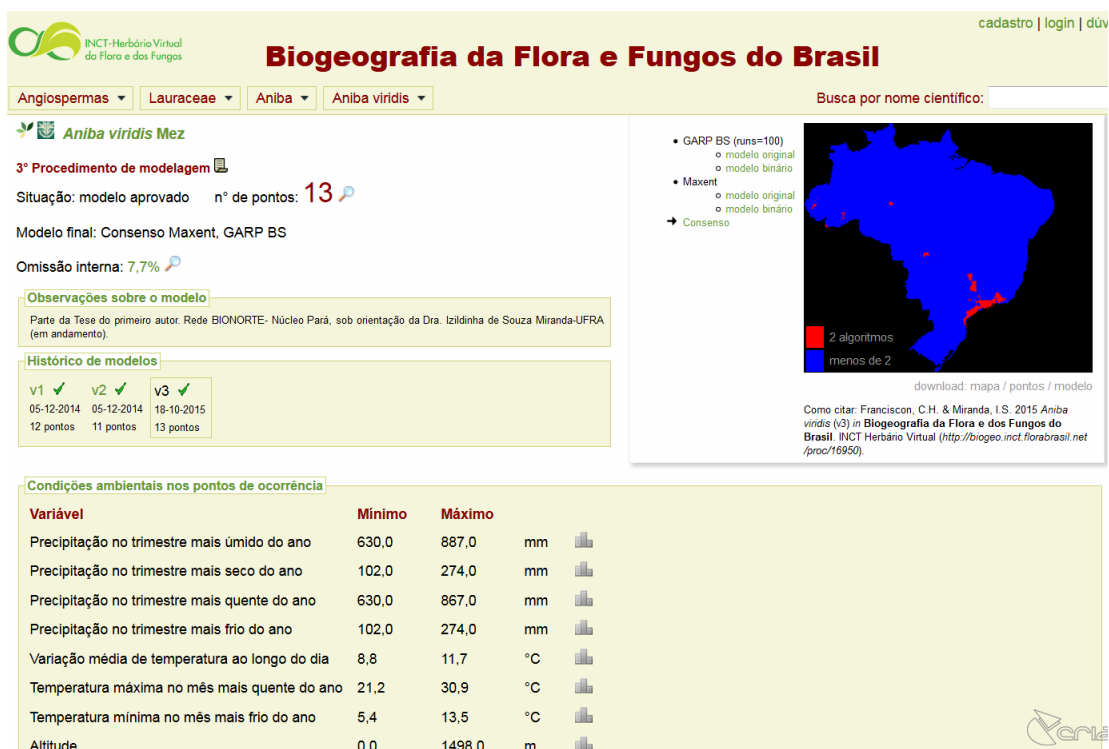
Figure 4. Example of a distribution model published on-line

Methodology

When the survey was sent out, there were 177 specialists registered in the system, meaning all were authorized to produce distribution models. Nevertheless, only 43 had actually published models on-line.

An email was sent to all 177 specialists who were asked to fill out the following information:

(1) Institution
(2) Academic level
(3) Field of interest
(4) If registered and did not publish any model, what was the impediment
(5) Use of the model published through Biogeo
    a. Planning new collecting efforts
    b. Public policies
    c. Articles
    d. Others (please specify)
(6) Weaknesses of BioGeo
(7) Strengths of BioGeo
(8) What would you like to see in BioGeo in the future

Results

We received 17 answers from 10 PhDs, 4 students doing their PhD, 2 Masters and 1 Bachelor (biology).

Nine (53%) did not publish their models for diverse reasons:

- They are still compiling the data
- They just carried out some tests
- They do not like the idea of having another specialist authorizing their participation – NOTE: this is not correct and an email was sent to this specialist to clarify this point.
- Problems with understanding the system
- The resulting model did not answer the hypothesis raised about the species
- Lack of time
- I am not a specialist

As to the use of the models:
- Two (12%) have not used any model available
- Four (24%) have used the models to plan new collecting efforts
- Three (18%) for public policies
- Seven (41%) to publish articles
- Two (12%) indicated using the models for their research (this was under "others")

Weaknesses
- The system does not include data from other South American countries (restricted to Brazil)
- The fact that non specialists can generate models
- Many specialists in biogeography are not participating. Suggests a better communication strategy to make BioGeo known.
- Data of the models can only be exported in xml. It would be interesting to have other formats such as ascgrid and float.
- The models use a fixed set of environmental layers. It is not possible to select layers, algorithms, and other parameters.
- Not intuitive
- Insufficient data, this is not a problem of the system but it affects the quality of the models generated
- Not being able to project future scenarios with climate change

Strengths
- Easy to use (5 – 29%)
- Available distribution models
- The system that generates the maps is very good
- The system is fast and the graphic output of the models is good
- Extremely easy to use, principally for decision making, policy or research
- Many species have models
- Data sharing without restrictions
- Standardization, enabling the comparison of models
- Treatment of the occurrence points to generate the models
- Reduction of processing capacity of desktops to generate models
- Utility in planning new collecting efforts
- Visualizing the models, data used, liberty to select the data (validate or not and justify)
- Facility to manipulate and interpret

- Potential for diverse uses, both academic and for public policies

Future Requirements
- Inclusion of data from other South American countries
- Historical data about areas where specimens were collected
- Specialists that participate in the Flora of Brazil initiative should be invited to participate
- A link to data use, restrictions, models and results to give proper credits and stimulate new contributions
- A feedback mechanism for public policies
- Distance training in GEOstatistics
- Possibility of selecting geographic areas (such as states, regions, phytogeographic limits, among others) and bioclimatic layers.
- Where specialists are listed, include the species and families they are modeling
- Enable projection for scenarios of climate change

Comments

Although the number of answers was small (less than 10% of those that produced models), there are important contributions that can help guide future activities.

The number of people that registered compared to the number of specialists that are actually producing models indicate that there may have been a misinterpretation as to the usage of the system. It seems that people may have thought it necessary to register in order to access the models. This seems plausible when some of the answers received by those that have not published any model includes "I am not a specialist".

One of the motivations in participating is obviously to use the model that was produced. Our main focus as to data users is the scientific community. Therefore, it is interesting to recognize the usage of the model to plan new collecting efforts, for research, to publish articles, and for public policies. These were all aims of this tool.

As to weaknesses, as the focus of the system was to help plan new surveys in Brazil, the geographic scope was Brazil and with current environmental conditions. It would be interesting to further develop the system increasing the geographic scope to South America – more data and possibly a better resulting model – and to build scenarios considering climate change.

A weakness mentioned referred to the necessity of a better communication strategy, as important specialists are not involved. This is true. The system was not publicized as it was under development and now, as the project ended, it is just being maintained. If we are able to obtain new grants, a communication strategy must be in place.

As to strengths, although one user said the system is not intuitive, five said it is easy to use. Given the fact that the system shares distribution models for 3.562 species without having provided any training courses, one can assume that it is intuitive for a knowledgeable person. Table 2 shows the number of species with models per taxonomic group.

Table 2. Species distribution models produced through BioGeo

Taxonomic Group	No. of Species in the List of Brazil	Species with Distribution Models	%	No. of Approved Models	No. of discarded models	Models awaiting approval
Algae	4.747		0			
Angiosperms	32.831	3,471	10.6%	4,046	126	147
Bryophytes	1.524	5	0.3%	5	3	12
Fungi	5.712	10	0.2%	10	6	1
Gymnosperms	30	4	13.3%	4		
Ferns and Lycophytes	1.253	59	4.7%	68	5	7
Total	46.097	3,549	7.7%	4,133	140	167

Despite the lack of a more substantial dissemination of BioGeo, almost 8% of all fungi and plant species that occur in Brazil have a distribution model published and openly available on BioGeo. It is interesting to see that one user expressed as strength of the system the fact that data is shared without restrictions. The answer as to future requirements certainly will help us when writing a new proposal for funds to enable the continuity of this initiative.

These analysis that involve direct contact with those that are contributing with data (herbaria) or with their knowledge (Annotation, BioGeo) are normally overlooked by e-infrastructure that are totally focused on the outputs of the project, both quantitative and qualitative indicators, and this way miss analyzing possible outcomes. This project brings another dimension to the work that is being carried out.

4. Project Implementation and Management

Completed Activities (February 2015 - 2016) – in bold activities that had not been foreseen	
Analysis of Data Providers (herbaria)	*Classification of herbaria, interview & preparation of questionnaires; application of the questionnaires; tabulation of results and preparation of the report; dissemination.* **A face-to-face meeting was held, organized by the Virtual Herbarium Steering Committee**
Analysis of Contributors (Annotation & BioGeo)	Preparation and Application of the questionnaire; Analysis and report; Dissemination
Data/System Users	Analysis of usage & defining a strategy for year 2

Assessment of the role of the virtual herbarium for graduate courses. **This activity planned to be concluded in year 2 was developed in year 1 as a face-to-face meeting was held and discussed the questionnaire and the SWOT analysis.**	Discussion of a strategy with the Virtual Herbarium Steering Committee. **As a preparation of a face-to-face meeting a SWOT analysis was sent together with the questionnaire and was discussed at the meeting**

Pending Activities (March 2016 - February 2017)

Data/System Users	*Developing an on-line survey to obtain feedback from users as to what the data is being used for and what is the profile of the users.*
Assessment of sensitive data	Evaluate what data is being blocked and for what purpose.

5. Project Outputs and Dissemination

Workshops held: February 2015 - 2016

Name of workshop	Outcome(s) of workshop	Number of participants present	Any relevant links to event information
Strategic Planning meeting of Brazil's Virtual Herbarium	Evaluation of the questionnaire (project outcomes) and SWOT analysis	50	This was a back to back meeting with an international symposium on Collaborative Networks and Digital Documentation about Biodiversity (http://www.cria.org.br/eventos/inct_i3b/index)

Conferences Attended (to discuss Open Science)

Name of Conference	Your contribution to the event	Outcomes of the conference? (collaborations, contributions, etc.)	Any relevant links to event information
Conference IUBS 2015 frontiers in Unified Biology and 32nd IUBS General Assembly Round table: *Unifying Biology through informatics*	*Challenges in building an infrastructure for all of biology: The importance of local infrastructures*	Dissemination of outcomes from data sharing (OCSDNet project)	

| IDigBio Summit 2015. Arlington, Virginia, USA. November 4-6, 2015 | Brazil's Virtual Herbarium: Outputs, Outcomes, and Challenges | Exchange of experiences with iDigBio team. | https://www.idigbio.org/wiki/images/9/96/IDigBio-Summit-V_Brazil-Virtual-Herbarium_Canhos.pdf |

Partnerships Formed to date:

Name of Partner	Type of stakeholder (policy maker, research organisation, community group, etc.)	Briefly describe your collaboration with this partner
OCSDNet team and associated projects	Community group	Although our collaboration is limited, as our team is small and our main concern is to develop all activities of the present project, I consider the interaction with OCSDNet team and members of the other OCSDNet projects an opportunity to expand our concept of open science.

6. Impact

The work carried out within the context of this project in 2015 produced a great impact on the Brazilian herbaria that are sharing their data through *Brazil's Virtual Herbarium*. The preparation and face-to-face discussions led to a change in thinking, not only on individual problems and outcomes but also on collective impacts and results. Together, the questionnaire and the SWOT analysis were fundamental in promoting a discussion on the impact of data sharing from the data providers' perspective and on the strengths, weaknesses, opportunities and threats concerning the Virtual Herbarium, as a basis for planning its future. This discussion helped individual herbaria realize their importance and role in the network. Another important point is the realization of the strength of the network, which is easily overseen on a day-to-day basis. At the same time that individual herbaria are seen as an essential element of the network, the achievements of the network, and here the IT team (CRIA) and steering committee are included, are understood as something only obtainable when there are collective strategies and goals in place.

This research project shows the importance of the e-infrastructure, not only to integrate data from different herbaria, enable data sharing, monitor the progress of the project, and offer a number of tools, but also to integrate special interest communities, giving means for social networks to innovate and expand.

Besides collaborating with the advancement of botany, the fact that all data is openly available means that it is also available to all communities from different fields of interest. During 2016, within the context of this project, the user community will be object of research, identifying the uses of data and the profile of the users.

The will to share data and information is important but not sufficient. This project makes an important contribution to open science, showing that data sharing must be planned, organized, funded, and community driven. Internationally accepted protocols and standards must be used, especially in the case of biodiversity, an important topic for one of the world's global challenges – sustainable development. Open science depends on open data and, with the developments in information and communication technologies, on open processes.

Brazil's Virtual Herbarium:

- Shows an e-infrastructure that aggregates data from different providers that continue to retain full control over the data that they share;
- Presents a number of on-line applications and tools, that enable different analysis to all interested;
- Presents indicators that allow measuring and visualizing the evolution of the e-infrastructure; and,
- Presents tools that enable an exchange of knowledge and information between data users and data providers.

As to measuring and evaluating the impact of Brazil's Virtual Herbarium, this project in 2015 measured its impact in the participating herbaria. The survey indicated that due to their participation in the project, the herbaria achieved greater institutional recognition (82%), greater involvement with graduate courses (67%), increase of the number of visits to the herbaria (85%), increase of its holdings (77%), and increase in external grants (51%).

We believe that being part of OCSDNet represents an opportunity to think beyond the project, understanding differences in the many fields of knowledge and communities. We also believe that this experience, with an important component of information and communication technology, may contribute to the thinking of other projects.

7. Reflective Learning

Since Brazil's Virtual Herbarium and CRIA itself, have open data as a value and an aim, participating in this project made us think further, to evaluate the outcome of data sharing to data providers. Before we were acting to fulfill our role, in developing the e-infrastructure and providing support to data providers and users. The application of the questionnaire and SWOT analysis and its discussion in a face-to-face meeting promoted an important change that is to think and act collectively. Interactions become more frequent and important demands and new developments are not just presented, but discussed.

The importance of our participation in OCSDNet, outside the scope of Brazil's Virtual Herbarium, was only realized by our team at the OCSDNet meeting in Bangkok. This is seen as an opportunity to learn different aspects of openness, but also as a challenge for the team, that has the contracted project to carry out, together with many other demands.

8. Recommendations (for OCSDNet)

We are focused on carrying out our project and have not been able to follow all the communication flow that OCSDNet is promoting. It would be interesting to have a feedback on whether our project is the only one facing this problem or if other projects are also not being able to keep up with the expected interaction with OCSDNet.