

## MID-TERM ACTIVITY REPORT

BIFA3\_47 - Georeferencing and Mobilization of Plant Species occurrence data from Pakistan

*Guidelines on how to complete the activity report are included in italics.*

### Contents

Executive summary .....	1
Contact information .....	1
Introduction.....	2
The project and its objectives.....	3
Activities .....	4
Deliverables.....	7
Calendar of activities .....	8
Project communications .....	10
Mid-term evaluation findings and recommendations for the remaining project implementation period .....	11
Annex – Sources of verification.....	13

---

### Executive summary

*Provide a brief explanation of the context and the **approach taken for the mid-term evaluation, and a summary of the main conclusions, lessons learned and recommendations for the remaining project period.***

Our project’s primary goal is mobilization and georeferencing of specimens from Pakistani herbaria, particularly those of Quaid-i-Azam University and the University of Swat. Another major goal is to develop current information on the number and status of all Pakistani herbaria and provide it to both Index herbariorum and GBIF, and to engage additional herbaria in data mobilization. So far the emphasis has been on mobilization of data at ISL and SWAT. The last goal is preparation of a data paper based on the Flora of Pakistan and information for District Swat. Progress has been evaluated primarily by evaluating relevant numbers. Lubna (project manager at the Quaid-i-Azam University) and Zahid Ullah (project lead at University of Swat) identified some significant impediments to progress, primarily difficulty of working over the web but also failure to adequately account for other factors affecting productivity in Pakistani herbaria. Steps for accelerating data capture have been taken. Another disappointment has been the number of individuals who accepted invitations to participate in the first workshop but simply failed to show up. The next workshop, which is currently being planned, will be restricted to the primary participants because it will focus on georeferencing of existing data.

---

### Contact information

Please provide the name, institutional affiliation, role in the project and contact details of the author(s) of the report

Dr. Mushtaq Ahmad, Primary Project Lead.

Herbarium of Pakistan (ISL)  
Department of Plant Sciences  
Quaid-I-Azam University  
Islamabad, Islamabad 45320  
Pakistan.  
Email: [mushtaq@qau.edu.pk](mailto:mushtaq@qau.edu.pk); [mushtaqflora@hotmail.com](mailto:mushtaqflora@hotmail.com)

Lubna, Digitization Supervisor at Quaid-I-Azam University  
Herbarium of Pakistan (ISL)  
Department of Plant Sciences  
Quaid-I-Azam University  
Islamabad, Islamabad 45320  
Pakistan.  
Email: [lubnafareed91@gmail.com](mailto:lubnafareed91@gmail.com)

Dr. Zahid Ullah, Project Lead at University of Swat

Centre for Plant Sciences and Biodiversity  
**University of Swat Pakistan**  
Sector-D Kanju Town 19201  
Swat, Khyber Pakhtunkhwa  
Pakistan  
Email: [zahidtaxon@uswat.edu.pk](mailto:zahidtaxon@uswat.edu.pk)

Dr. Neil Cobb, Project Lead at Northern Arizona University

Merriam-Powell Center for Environmental Research  
Northern Arizona University  
805 S. Beaver Street  
PO Box: 6077  
Flagstaff, Arizona 86011-6077  
United States  
Email: [Neil.Cobb@nau.edu](mailto:Neil.Cobb@nau.edu)

Dr. Mary E. Barkworth, Project Advisor and Report Author

Intermountain Herbarium  
Utah State University  
Logan, Utah 84322-5305  
United States  
Email: [mary.barkworth2usu.edu](mailto:mary.barkworth2usu.edu)

---

## Introduction

- *A description of how the evaluation has been carried out (e.g. consultation or surveys with project partners and participants). Please refer to the description of monitoring and evaluation plans in the original project proposal.*
- *A description of how the project partners will use the evaluation results.*

Goal 1. Mobilize 37,000 records, 36,000 from the two participating herbaria and 1,000 from other Pakistani herbaria.

Workshops are evaluated by asking participants to respond with suggestions for improvements and monitoring impacts.

Data capture and mobilization progress is evaluated by a) monitoring records made available via OpenHerbarium and b) progress in becoming data providers to GBIF.

Follow up has been via Email and involved creation of a spreadsheet for use in data capture and export. In addition, software is in development that will benefit data

capture from existing specimens and generation of high quality labels and easy data ingestion for new collections.

Goal 2. Georeference 37,000 mobilized records.

Progress will be monitored and evaluated by monitoring records in OpenHerbarium. The activity has not yet been started (this is in accordance with the proposal). Barkworth is exploring approaches

Goal 3. Engage at least 10 taxonomists not employed on the project in data mobilization.

Progress is evaluated by monitoring inquiries, watching addition of records to non-funded herbaria, and use of tools for benefiting from resources associated with Pakistani data in OpenHerbarium and KeyBase.

Drs Abdur Rashid (University of Peshawar) and Rabia Memon (University of Sindh-Jamshoro) are arranging another workshop on digitization that will also involve representatives of the University of Karachi and initiate preliminary work on mobilizing data from macro algae.

Goal 4. Develop information on herbaria in Pakistan.

Progress will be monitored by returns to survey and watching entries on Index herbariorum.

No herbaria have responded to the survey. We shall try personal contact with the collection managers or their departments.

Goal 5. Data papers on a) Flora of Pakistan and Flora of District Swat.

Progress will be monitored via email contacts and by recording preparation of drafts. This activity has not yet been initiated.

Barkworth has begun reviewing the data from the flora that are available online via OpenHerbarium. This has revealed many gaps in the data that she is now addressing.

---

## The project and its objectives

*A brief summary of the project to help readers understand its objectives, including, for example:*

- *The project's start date and expected duration*
- *A list of project participants and description of the main stakeholders*
- *The targeted capacity needs as outlined in the project proposal*

The project is designed to help embed data sharing into the activities of botanists in Pakistan. Its specific goals are to mobilize and georeference 37,000 records from Pakistani herbaria, engage at least 10 taxonomists outside the funded herbaria in some aspect of data sharing, update and expand existing information on herbaria in Pakistan, and prepare two data papers relating to Pakistan's flora. Project participants

The project began on April 1, 2018 and has a duration of one year.

### Project partners

Dr. Mushtaq Ahmad and Lubna, Herbarium of Quaid-i-Azam University (ISL)

Dr. Zahid Ullah, Herbarium of University of Swat (SWAT)

Neil Cobb and Mary Barkworth, Northern Arizona University

### Other named participants

Herbarium and Botanic Garden, University of Balochistan

Herbarium, Pir Mehr Ali Shah Arid Agriculture University

## Activities

Please indicate the status of the activities as outlined in the project proposal. The table below should be completed in the same way as in the full proposal but should include information about the status of the activity.

Description of activity	Partners involved	Contribution of activity to goals listed in table 4.3	Status of activity as of mid-term reporting
<b>Digitizing and publishing georeferenced species occurrence data based on specimens held in Asian collections</b>			
Develop pdfs on fields used for capturing data and georeferencing.	NAU	<b>Goal 1:</b> helping personnel provide high quality data in an appropriate manner.	Completed; subsequently developed spreadsheet to facilitate data capture.
Hold participant's workshop on data capture at ISL	NAU, ISL, SWAT, SBKWU	<b>Goal 1:</b> help personnel understand the workflows involved and the purpose of each step.	The workshop was held but the only herbaria represented were SWAT, ISL, RAW and SBKWU the herbarium of the National Agricultural Research Council. Individuals from two other Pakistani institutions (Islamia College Peshawar and the University of Peshawar) assisted Barkworth in the presentation. Both had assisted in two other workshops conducted elsewhere prior to the start date of the project workshop.

<p>Develop offline database for data capture</p>	<p>NAU</p>	<p><b>Goal 1:</b> enable quality data entry offline that incorporates basic tools for ensuring accuracy in entering scientific names and administrative regions.</p>	<p>A program, SPINDLE, is being developed at Utah State University (using funds provided to NAU) that allows for data capture, validation, and export offline. It can be used both in the field for new collections and in the herbarium for existing collections. The data capture portion has been developed; data validation and label printing will be added before the end of the year. A working prototype is available online through GitHub (<a href="https://github.com/sylviakinosiain/spindle">https://github.com/sylviakinosiain/spindle</a>). The program, in its current state, uses the JavaScript engine embedded web browsers released within the last decade to validate herbarium label data, and output a comma-separated value (CSV) file of the user's data. While the program is not yet complete, it has basic data-validation and export capability, and it is our intention to begin a formal launch in ~December 2018 once it has been tested by users both in Pakistan and the US.</p>
<p>Hold workshop on georeferencing, including batch georeferencing</p>	<p>NAU, ISL, SWAT</p>	<p><b>Goal 2:</b> Familiarize participants with georeferencing fields.</p>	<p>The workshop is in the process of being scheduled. It will be held in late November. In the meantime, Barkworth is exploring procedures that are less reliant on internet access than those included in the proposal. A first step is determining from the data already entered how many different locations are represented.</p>
<p>Georeference unique localities from Flora of Pakistan</p>	<p>NAU, SWAT, ISL, QUETTA</p>	<p><b>Goal 2:</b> Georeference 37,000 records. Accelerate look up by creating a file for frequently visited localities.</p>	<p>The data acquired through the procedures described above will be used to develop a georeferenced file of frequently visited localities.</p>

<b>Compiling inventories of biodiversity data holdings (for example, by implementing metadata catalogues)</b>			
<p>Write to all Pakistani herbaria listed in Index herbariorum asking for updated information on both specimen holdings and data records. Follow up with phone calls, possibly to the institution, if no response is received.</p>	<p>ISL, SWAT</p>	<p><b>Goal 4:</b> Compile inventory of specimens and data records currently associated with Pakistani herbaria. Will determine by taxonomic categories used by the Index.</p>	<p>A survey form was sent to the herbaria listed in Index herbariorum in July. None have responded. The process of calling individual herbaria will begin in November.</p>
<b>Preparing <a href="#">data papers</a></b>			
<p>Review data, family by family, from Flora of Pakistan, revising nomenclature where appropriate</p>	<p>ISL, SWAT, NAU and invite others</p>	<p><b>Goal 5.</b> Data paper on Flora of Pakistan. Note changes.</p>	<p>This has been started. Its clear that data from several volumes had not been uploaded prior to the start of this project. It is now being added but the data paper may have to describe the task as more incomplete than anticipated.</p>
<p>Compare species from Pakistan checklists with those from Flora</p>	<p>ISL, SWAT, NAU and invite others</p>	<p><b>Goal 5.</b> Data paper on Flora of Pakistan. Note changes to be made.</p>	<p>The comparison between the records from ISL and SWAT with the Flora of Pakistan has begun. It was discovered that the entries in OpenHerbarium from the Flora of Pakistan are not complete, even for volumes prior to 220. So far, once nomenclatural and taxonomic changes have been taken care of, no new taxa have been added. A comparison with records from GBIF has not yet been made.</p>
<p>Compare species in GBIF with those online checklists and determine which reflect differences in taxa reported, which difference in nomenclature;</p>	<p>ISL, SWAT, NAU</p>	<p><b>Goal 5:</b> Show how useful mobilized data can be.</p>	<p>The comparison between the records from ISL and SWAT with the Flora of Pakistan has begun. It was discovered that the entries in OpenHerbarium from the Flora of Pakistan are not complete, even for volumes prior to 220. So far, once nomenclatural and taxonomic changes have been taken care of, no new taxa have been added. A comparison with records from GBIF has not yet been made.</p>
<p>Flora of District Swat</p>	<p>SWAT</p>	<p><b>Goal 5:</b> Goal: demonstrate value of mobilizing and sharing data.</p>	<p>Work on this paper has not yet begun.</p>

Offer to work with Pakistanis who have published checklists to make them available as data papers	NAU, SWAT, ISL	<b>Goal 5:</b> Increase knowledge of existing data resources and expand engagement.	This offer will be made in November and repeated in March. An alternative suggestion will be made, that individuals publishing checklists make them available as checklists on OpenHerbarium, thereby making their results easier for other to find, easier to download, and enabling them to be viewed with images, to the extent that these have been added to OpenHerbarium.
<b>Other activity types</b>			
Hold blended workshop on value of engaging in data mobilization for individuals and collections (subject to receipt of funding from Higher Education Commission of Pakistan)	ISL, SWAT, NAU	<b>Goal 3:</b> Develop sustainability for data mobilization and sharing	The intent was to hold this late in 2018 but Barkworth's US visa requires that she spend at least 6 months and 1 day in the US each year. Consequently, only a short time is available for this activity in 2018. A funding proposal for a workshop at the University of Sindh Jamshoro on December 1-3 is being prepared. Barkworth has begun revision of materials needed for starting digitization of a herbarium for use at this workshop.

## Deliverables

*This section should summarize the project activities completed by the mid-term, with a description of the associated outputs and deliverables. Please highlight any changes from the original plans provided in the full project proposal.*

*If no result has been achieved on a specific point, please indicate it as "no result achieved yet".*

### a. Data

*Details of datasets expected to be mobilized as an outcome of the project:*

<b>Title of dataset</b>	<b>Taxonomic/geographic scope</b>	<b>Approximate number of records (specimens)</b>	<b>Current format (e.g. undigitized, digitized)</b>
ISL (Herbarium of Quaid-i-Azam)	Vascular plants/Pakistan	180,000; target is to database and georeference (D&G) 30,000 in this project	12000 records databased; 1 georeferenced.
SWAT (Herbarium of Univ. of Swat)	Vascular plants/Pakistan	6000	453 databased; 10 georeferenced (another 43 have

			coordinate data); 20 fungi also databased.
Flora of Pakistan	Vascular plants/Pakistan	60,000	32,185 databased; 31235 georeferenced to 2 degree grid, i.e., coordinate uncertainty 293 km. This represents approximately half the families.

### b. Other deliverables

*Describe other deliverables expected from the project (e.g. publication of data papers, catalogues, reports etc.)*

Data papers:

Flora of Pakistan (based on print version, comparison with contributed records).

Flora of Swat (evaluation of impact of collections from SWAT and ISL on estimates based on Flora of Pakistan)

Explanatory materials

Multiple documents on use of Symbiota

Documents relating to the basic principles of data mobilization, sharing, georeferencing, visualization, and analysis. All will be written at an introductory level and include questions to help students master the material treated.

SPINDLE, a program for recording and validating data offline, generating specimen labels (for newly collected specimens), and exporting georeferenced occurrence data.

### Calendar of activities

*The calendar should be completed in the same way as in the Full Project Proposal (4.6) but should include any expected changes. Please provide reasons for any expected changes in the Notes column in the table below.*

Proposed dates	Activity	Lead partner	Notes
June 2018	Attendance of project team member at BIFA Capacity Enhancement Workshop		Lubna (ISL) attended because Dr. Mushtaq Ahmad (ISL) was engaged in the other assignments of the project.
April 1	Initiation of data mobilization using available data capture tools; new or modified tools will be modified as they become available. Barkworth (NAU) had already planned to be in Pakistan in March and April; she will devote part of April to this workshop.	ISL	Held; only one non-project herbarium attended although others had accepted an invitation to participate.



April 1	Initiate steps for becoming GBIF data providers	NAU to coordinate exposure of records but each herbarium will be a separate provider.	
Each month	Monitor mobilization progress weekly, summarize monthly. Identify impediments and modify procedures to achieve faster rate as tools become available.	NAU (Barkworth)	Monitoring resulted in a switch to use of spreadsheets for data capture. Lubna is uploading data from ISL; Barkworth is uploading data from SWAT. Software for enabling validated data capture offline and label preparation is in development.  Barkworth has approached the herbarium directors at the Universities of Sindh-Jamshoro and Peshawar about initiating databasing at their two institutions.
April 1	Start contacting herbaria concerning their holdings. Summarize results in June. Identify reasons for failing to register with or update Index herbarium.	SWAT	Survey sent out in July. No responses have been received. Will consider other approaches to obtaining the data in November.
July	Presentation on data mobilization in Pakistan at Botanical Society of America (not included in budget; Barkworth (NAU) will be attending regardless of funding)	NAU (Barkworth)	Paper presented. OpenHerbarium web site modified to give greater prominence to Pakistani data sources. Contract given to Sylvia Kinosian, using IT matching funds, to develop offline app for data capture.
August	Presentation on modifying software for use in countries with poor internet access to TDWG (not included in budget; Barkworth (NAU) will be attending regardless of funding)	NAU (Barkworth)	Paper presented. ISL became a GBIF data provider, an achievement that was reported on the Facebook page of the Pakistan Botanical Society.

Sept.			SWAT, using other funding, hired individuals to help process specimens, a necessary pre-digitization step.
Oct.			Presentation on biodiversity mobilization scheduled for iDigBio summit but displaced by extended discussion of NSF-funded project, Symbiota2, which will include modifications designed to make it easier for collections and collectors in countries such as Pakistan to share GBIF-worthy data.
November	Blended workshop promoting data mobilization (subject to receipt of funding from funding agency in Pakistan)	ISL, NAU (Barkworth)	Being scheduled in conjunction with another workshop in basic data mobilization at the University of Sindh Jamshoro.
November	Submit Flora of District Swat as data paper with online, illustrated checklist	SWAT	Will be later than this.
December	Submit Flora of Pakistan data paper	NAU	Will be later than this.
December	Submit abstracts, and encourage others to do so, on their data mobilization activities to annual meeting of Pakistan Botanical Society.		
December	Identify potential sources for further work on digitization and proposal deadlines; work with others to ensure at least one more funding proposal is submitted in 2019.		
March	Prepare report on progress and impact.	NAU	

## Project communications

*Describe the plans to communicate and share the results of your project with the project stakeholders and broader GBIF community. Please also review the page describing your project available from*

<https://www.gbif.org/project/5AVwqL7izSysQO2mquiYQ8/georeferencing-and-mobilization-of-plant-occurrence-data-from-pakistan>. Highlight any additional documents, events, news items or links that you would like to add to your page.

We use the Facebook page of the Society for Pakistan Plant Taxonomists to communicate significant milestones to taxonomists within Pakistan. Barkworth has given two presentations that cite the project as the primary source of ideas. Dr. Will Pearse (Utah State University) gave a talk to Symbiota-based project leaders at the iDigBio Summit on Oct. 3 describing SPINDLE, a software program funded in part by this project. Proposals seeking additional funding will cite information and resources developed for and through the project.

---

## Mid-term evaluation findings and recommendations for the remaining project implementation period

*This should be the main section of the report, covering for example:*

- *An evaluation of the project activities by the mid-term and their outputs/deliverables*

Records to GBIF: ISL is on track to meet the targets; SWAT had to overcome some obstacles but anticipates being able to achieve the target by the end of the project. Other institutions: participation in actually providing records has been disappointing but there is continued interest, importantly from the University of Karachi, headquarters of the Flora of Pakistan and home to Pakistan's major macroalgae collection, and the University of Sindh-Jamshoro, with a collection that emphasizes one of Pakistan's poorly documented regions.

Georeferencing: not yet started; reliance on use of Geolocate via the internet from Pakistan is not feasible so other approaches are being examined. These will involve Pakistani participants providing data for key locations via a spread sheet. The approach will be finalized after consultation with Nelson Rios, designer of Geolocate, who has been working with similar issues from other groups. Uncertainties for many records will be large because record data is often imprecise.

Engage at least 10 non-participant taxonomists in data mobilization: No success so far but hope to change this in November.

Information on herbaria in Pakistan: Disappointing; no responses received. Will try to identify the problem and obtain information by the end of the project.

Data papers: As planned, these have not been started but it may be that their completion will extend beyond March 31 because our emphasis is on expanding the number of records being made available.

- *Any feedback on the project's relevance from the partners and stakeholders*

Barkworth (NAU): The project has succeeded in increasing interest in data mobilization within Pakistan. Prior to the award, Dr. Zahid Ullah was the only Pakistani plant taxonomist contributing records for sharing. It has also clarified, and enabled addressing, some of the impediments to data mobilization that Pakistani botanists face. Several of these impediments are the same as those faced by botanists in other countries so the solutions will extend beyond Pakistan, extended in part through Cobb and Barkworth's in participation in Symbiota2, a project funded by the US National Science Foundation.

- *Comments on the project implementation, its efficiency and effectiveness*

It would be helpful if the performance part of the project could begin after the award is announced. Alternatively, if it could be stated that the completion date (and payment of the final installment) could be made 15 months after the announcement date. The reason for this suggestion is that Swat was unable to advertise a position until the money for the position had been provided. That, plus the time required for advertising the position, meant that, in effect, it could not start until about 3 months after the start date.

One area where project implementation was more difficult than anticipated were the inability to use browser-based data entry effectively. This has been addressed, as mentioned earlier, by switching to use of a spreadsheet but the greater and more substantial benefit will come once development of SPINDLE, an app for capturing and validating biodiversity data in the field and herbaria, generating specimen labels, and creating csv files for export to Symbiota networks OFFLINE.

The second major problem arose from the situation at the University of Swat. On October 25, 2015, many of the university buildings, including those housing the herbarium, were destroyed in a magnitude 7.5 earthquake. Since that time, the herbarium has very limited space on campus. What was not clear was that, at the start of the project, many of the 6000 specimens mentioned in the proposal existed but had not been labeled and mounted. The combination has seriously impeded the proposed work flow at SWAT. To address the problem, Dr. Zahid Ullad has used other funding to purchase curatorial supplies and hire undergraduates to start processing the existing specimens. Some of the work is done in the dormitory where the students reside. He anticipates that this change, plus the switch to the use of spreadsheets for data capture, will enable SWAT to hit the target for data mobilization by the end of the project. The existing plans call for moving the herbarium to more appropriate, permanent space in 2019.

- *The management arrangements for the project, including support from the GBIF Secretariat*

We have been learning how to manage the project as we go. One difficulty is the difference in time zone; another has been a greater need for formal documentation by toh Quaid-i-Azam University and Northern Arizona University before money provided by GBIF could be transferred to NAU.

Barkworth found it very helpful to sit down with a GBIF representative at TDWG to review the comments generated by using GBIF's data validator.

*Any reflection on the mid-term evaluation itself that could help inform the project's final evaluation and final report*

Being required to prepare the evaluation has been useful but it does seem to ask for similar information in many different ways.

- *Areas of success to build on during the remainder of the project implementation*

Some data have been mobilized. Now we need to start encouraging its use. Related activities, not included in the proposal, will aid in this regard. They include providing online access to keys in the Flora of Pakistan via KeyBase and linking these to pages generated by OpenHerbarium. The taxonomy is being updated at the same time. In addition, images and descriptions are being added to OpenHerbarium. Once we start georeferencing information, the site will start providing integrated access to Pakistan-oriented floristic information and the value of sharing data will become more apparent. We have asked organizers of the annual meeting of the Pakistan Botanical Society for information regarding organization of a symposium on some aspect of data sharing as part of that event. We hope that GBIF will be willing to send a representative if the organizers accept our request.

Creation of SPINDLE will be a substantial benefit to people in many countries. The code, which is still in development, is available at <https://github.com/sylviakinosian/spindle>.

*Try to clearly document any changes to the project plans that will be made based on the findings of the mid-term evaluation. Please discuss any substantial changes with the GBIF Secretariat ([bifa@gbif.org](mailto:bifa@gbif.org)). In addition, please outline any recommendations for the GBIF Secretariat or the community to reinforce the initial successes of the project.*

---

## Annex – Sources of verification

Sources of verification are for example links to relevant digital documents, news/newsletters, brochures, copies of agreements with data holding institutions, workshop related documents, pictures, etc.

---

### Links

Facebook-Pakistan Society of Plant Taxonomists. Used for communication. The announcement that records from ISL were now available via GBIF resulted in many likes and congratulations.

Geolocate: website with programs for assisting in georeferencing natural history collections. <http://www.geo-locate.org/>.

KeyBase: A web site for displaying dichotomous keys and linking the terminal taxa to other pages, such as taxon pages generated by OpenHerbarium. <https://keybase.rbg.vic.gov.au/>

OpenHerbarium: a Symbiota web site to which data are submitted and from which they are made available to GBIF. <http://OPenHerbarium.org/>

SPINDLE: an app for capturing plant specimen data both from existing collections and new ones that will validate the data, generate a label, and create the csv file needed to upload to Symbiota web sites (such as OpenHerbarium). <https://github.com/sylviakinosian/spindle>

Symbiota: a suite of programs for mobilizing, sharing, and visualizing biodiversity data. <https://github.com/Symbiota/Symbiota>. Help files are at <http://symbiota.org/docs/symbiota-introduction/symbiota-help-pages/>.

Symbiota Help: A compilation of resources for doing things in a Symbiota network. Some are pdf files, some videos, and some are web pages. <http://symbiota.org/docs/symbiota-introduction/symbiota-help-pages/>