

## Digitizing and data basing of bee specimens in Thailand

---

**Programme:**BIFA

**Project ID:** BIFA5\_005

**Project lead organization:**Chulalongkorn University, Department of Biology

**Project implementation period:**1/7/2020 - 30/6/2021

**Report approved:** 19/8/2021

### Narrative Final report

---

#### Executive Summary

---

The initial purpose of this project is to mobilize at least 8,000 processed occurrence records of Thai bees deposited at the Chulalongkorn University Natural History Museum (CUNHM), Thailand to GBIF. The objectives were met with 10,414 verified records of bees from 4 families: Apidae (6,651), Colletidae (11), Halictidae (1854), and Megachilidae (1898). Bees are identified to 39 described genera and 155 species despite 217+ morphospecies identified. More than 26,000 images from 6,500+ specimens were captured. 8,597+ out of 10,414 occurrence records are georeferenced.

Overall, the project implemented is a success despite many delays during to the global pandemic. Important lesson learned from this project including (1) establishing a realistic goal (2) appropriate task allocation and synergy among staff members (3) frequent contacts with GBIF personnel is highly encouraged (4) scheduled meetings with GBIF management team will clarify many questions relating to the project. Post-project activities that will insure the project sustainability including (1) proposing another data mobilization project to GBIF to include specimens collected from Thai national parks and wildlife sanctuaries that were deposited at the Queen Sirikit Botanical Garden (QSBG) in Chiang Mai through our proposal BIFA6\_009 (2) developing a DNA barcoding of Thai bee using the current project information as a background info to collaborate with iBOL. (3) The current database will be updated every three months after the end of the project with new occurrence records should GBIF permit us to continue working.

#### Progress against milestones

---

##### **Has your project completed all planned activities?: No**

**Rationale:** (1) Digitization II: Photographing of specimens and associated labels

Initially, we proposed to captured 8,000 images from 8,000 bee specimens (1 image for 1 specimen); however, we decided to produce at least 4 image views, which require more time to complete all 8,000 specimens. Currently, we captured 26,000+ images from 6500+ specimens and will continue after the final report is submitted until we reach the target with 32,000+ images.

(2) Digitization IV: Attaching QR codes to specimens and link to database

QR numbers were generated for each occurrence record. These codes will be uploaded later in the dataset (currently at 10,414 records). Attaching actual QR information to the specimens will be carried out after the submission of the final report.

(3) Workshop

The recent surge of COVID-19 cases in Thailand, particularly in Bangkok, interfered with our initial plan to host a workshop in June 2021. We are planning to host a workshop again in August 2021.

##### **Has your project produced all deliverables?: Yes**

## Activity implementation summary

### Mandatory attendance @ BIFA workshop

Our technician, Mr. Pakorn Nalinrachatakarn, attended the mandatory workshop on the 20th-24th July 2020 and completed the training with Basic badge and certificate (see attached document).

### Digitization I: Sort, identify, validate specimen data for registration into GBIF

More than 10,414 occurrence records were sorted, identified and validated for the Thai bee specimens. This is 2,900+ more records beyond our initial pledge of 8000 specimens.

### Digitization II: Photographing of specimens and associated labels

Initially, we proposed to capture 8,000+ images from 8,000+ bee specimens (1 image for 1 specimen); however, we decided to capture at least 4 image views including the actual label, which consumed more time to complete. Currently we captured 26,000+ images from 6500+ specimens and will continue until we reach 8000+ specimens with 32,000 images. The number of images that can be viewed in GBIF at the present is 14,000+ images from 3,500+ specimens that were already uploaded.

### Digitization III: Mapping of species distribution

8,597+ out of 10,414 occurrence records are georeferenced. This is 600 more georeferenced records beyond the initial pledge of 8000.

### Digitization IV: Attaching QR codes to specimens and link to database

QR numbers were generated for each occurrence record. These codes will be uploaded later in the dataset. Attaching actual QR information to the specimens will be carried out after the submission of the final report. This delay initially stemmed from the pandemic situation and the lack of funding during the early days of the project, which a suitable equipment for printing the actual QR codes cannot be obtained on time. However, the QR numbers are already generated and the final step of attaching the QR codes to specimens should be completed in less than a month.

### Transcription of data to GBIF platform and upload processed records

Since the beginning stage of the project, we uploaded our occurrence data twice in major batches. First attempt was on the 28th January 2021 for 7900+ occurrence records and the second attempt on the 1st July 2021 for a cumulative of 10,414+ records. This is 2,900+ more records beyond our initial pledge of 8000.

The dataset contains occurrence records (10,414) of bees (Anthophila) from 4 families: Apidae (6,651), Colletidae (11), Halictidae (1854), and Megachilidae (1898). 10,388 records have been identified to a total of 39 described genera and 155 species: Apidae 16 genera and 90 species, Colletidae 2 genera and species, Halictidae 11 genera and 25 species, and Megachilidae 10 genera and 39 species. However, the total number of morphospecies identified is 217+ suggesting more describing of new species and revision are essential for the future.

### Submission of mid-term report

The mid-term report was submitted on the 28th of January and was approved later by the GBIF secretariat.

### Workshop

Workshop to showcase Thai bee database to students and researchers interested in pollination ecology, conservation, bee taxonomy, and biodiversity informatics that was initially schedule for June 2021 was postponed due to the recent surge of COVID-19 cases in Thailand, particularly in Bangkok. Nevertheless, we are planning to hold an actual or virtual workshop depending on the COVID-19 situation Thailand in August 2021.

## Completed activities

### **Activity name: Mandatory attendance @ BIFA**

---

**Description:** Mandatory attendance @ BIFA workshop

**Start Date - End Date:** 20/7/2020 - 24/7/2020

**Verification Sources:** DOI10.15468/tf4ejd

---

**Activity name: Digitization I**

---

**Description:** Sort, identify, validate specimen data for registration into GBIF

**Start Date - End Date:** 1/7/2020 - 30/6/2021

**Verification Sources:** DOI10.15468/tf4ejd

---

**Activity name: Digitization III**

---

**Description:** Mapping of species distribution

**Start Date - End Date:** 1/7/2020 - 30/6/2021

**Verification Sources:** DOI10.15468/tf4ejd

---

**Activity name: Deliveries of Data**

---

**Description:** Transcription of data to GBIF platform and uploaded processed records

**Start Date - End Date:** 1/12/2020 - 30/6/2021

**Verification Sources:** DOI10.15468/tf4ejd

---

**Activity name: Submission of mid-term report**

---

**Description:** Submission of mid-term report

**Start Date - End Date:** 1/12/2020 - 31/1/2021

**Verification Sources:** Please check Approved Reports section.

---

**Report on Deliverables**

---

**Production of Deliverables - Summary**

The expected deliverable product of this project is a fully digitized occurrence dataset of 8,000 bee specimens in Thailand covering bees from families Apidae, Colletidae, Halictidae, and Megachilidae. Fortunately, we managed to process 10,414 verified bee occurrence records and uploaded the dataset in GBIF under DOI10.15468/tf4ejd. However, images were captured from 6,500 species with each specimens comprising 4 images each (26,000 images in total). After the submission of final report, we will continue producing images for the rest of 1,500 bee specimens to reach a total images of 32,000 pictures. The cause of delay for image capturing in the later half of the project is due primarily to the pandemic situation and cases that surged in Thailand which resulted in major lockdown at university where the project has been carried out.

**Production of deliverables**

---

**Title: Bees of Thailand**

---

**Type:** Dataset

**Status update:** 10,414 occurrence records with verified geo-referencing coordinates are completed with 26,000 images from 6,500 bee specimens.

**Dataset scope:** Bees in families Apidae, Colletidae, Halictidae, and Megachilidae in Thailand}

**Expected number of records:** 10414

**Data holder:** Chulalongkorn University

**Data host institution:** Chulalongkorn University

**Sampling method:**

**% complete:** 80

**DOI:** DOI

**Expected date of publication:**

---

**Impact of COVID-19 pandemic on project implementation**

---

The project has been conducted consistently from the beginning in July 2020; however, intermittent university shut-downs from COVID-19 situation has delayed our works dramatically after the mid-term report submission in January 2021. The shut-down of the university prohibited us from gaining access to collection, photographic equipment, and computer including server for database storage. Nevertheless, we were able to compensate the time lost and keep up with our work to come up with 10,414 occurrence records of bees specimens in Thailand, though the images captures came up short for 1,500 specimens. In addition, we encountered difficulty in obtaining suitable equipment and computer to initiate the project because trades were disrupted from the pandemic. We have to wait for suitable equipment to arrive before we can proceed with our work. This is the case for obtaining appropriate QR code printer and other computer components to link to the dataset.

In addition, workshop to showcase Thai bee database to students and researchers interested in pollination ecology, conservation, bee taxonomy, and biodiversity informatics that was initially schedule for June 2021 was postponed due to the recent surge of COVID-19 cases. Nevertheless, we are planning to hold an actual or virtual workshop depending on the COVID-19 situation in Thailand in August 2021.

In conclusion, despite the obstacles we faced during this pandemic, we are delighted that we achieve many milestones in our project to meet this deadline.

## Communications and visibility

---

Occurrence record database of bees in Thailand is link to the website relating to biodiversity information in Thailand hosted by the NSTDA (underdevelopment in <http://61.91.64.2:8075>). Taxonomically important records of certain species resulted from this work will be published in scientific journals to complement the database in the future. The processes and methods of digitization of bee occurrence records will be disseminated and shared with other research collections, universities, and institutions in Thailand through communications to the NSTDA website and other social media.

In August 2021, a workshop on Thai bee databasing is planned at Chulalongkorn University, Bangkok, which will bring together students and researchers interested in pollination ecology, conservation, and bee taxonomy to familiarize and learn how to efficiently utilize the information provided in the database, and also to advertise information regarding the importance of pollinators to the public. For further communication and dissemination of the project, we are more than happy to liaise with the GBIF Secretariat to help advertising of biodiversity data sharing to other Thai research institutions. In addition, working with the GBIF Secretariat in the future will surely provide opportunity for us to gain future insights and update our knowledge in biodiversity information management.

## Monitoring and evaluation

---

### Final Evaluation

Overall, the project implemented is a success with 10,414 occurrence records for bee specimens collected in Thailand from 1903-2021 published in GBIF with more than 26,000 high-resolution images. We are enthusiastic about this first major bee dataset for Southeast Asia to be available to the public and the scientific community. The number of specimen records that we initially proposed (8,000) is a practical number of record that can be achieved in a year worth of effort. Despite the delay due the global pandemic situation, we believe we are fortunate and consistent in executing the project to reach this current status.

The strength of this project implementation is the synergistic efforts among everyone involved in the database built up. Task allocation is very effective by assigning the right person to the right job. The recruitment of technicians working in the project with extensive background in biodiversity, photographing and computer is critical to the success of the work. In addition, the staff meetings were held very week instead of three month period as initially proposed. All of the staff worked in the same room and most of the time at the same time. This ambience created opportunity to address questions and discussion among technicians working on different aspects of the project.

Weaknesses observed throughout the project implementation period include factors involving shortage of computer and photographing equipment to process high quality images fast enough to meet the deadline. Further more, working on the project during the global pandemic affected our working schedule drastically, particularly after January to June 2021. The university where we worked was locked down periodically and the water, electricity, and the internet connection were frequently shut down. These internal problems directly and indirectly delayed our progress in finishing images capturing of 8,000 specimens.

One major change in the project original plan is instead of capturing one image per bee specimens, we decided that the dataset will be improved greatly by providing more views of the specimens. Therefore, we provides at least three image views and a picture of the original label for each specimen. This alteration tripled our work time on image capturing step and delayed our delivery of all images from 8,000 specimens. Nevertheless, we only have 1,500 specimens left to capture.

The GBIF Secretariat and the management team have been very helpful throughout the implementation of the project. Email responding were answered in short time with practical info to help us tackled many technical issues. In addition, the virtual meetings are informative and essential to our understanding of GBIF goals. All personnel working for GBIF are empathetic with good will.

The only issue that we would like to convey to the GBIF management team revolves around the mode of funding transfer. The bureaucratic process in Thai university hindered the efficiency for researchers to obtain monetary supports in time to meet many of their objective goals, i.e., the first installment of this project funding was released by GBIF to the university in July 2020 but the money was not release to the team until the end of January 2021. This situation caused delays on the project because many necessary equipment can not be procured. We are hoping that in the future, the BIFA grant will consider a direct transfer of funding to the project manager instead.

### Best Practices and Lessons Learned

Lesson learned from this project can be described as follows:

- (1) To accomplish a project relating to data mobilization and database construction, a realistic goal, i.e., number of records needed to be carefully and diligently plan. We were fortunate to complete our pledge in term of number of record because we did not overcommit in the beginning.
- (2) Task allocation is one of the major components in the success of a group project. In addition, consistent and swift communication among staff can reduce number of difficulties that may have occurred. We are grateful with the working atmosphere that we have on this project. Every staff is dedicate to his/her specialized task and the work ethic was held.
- (3) A consistent communication with GBIF management team whether scheduled or casual contacted provide many opportunities to discuss the progress, problem, and solution to project development throughout the year.
- (4) Mandatory and voluntary meetings scheduled by the GBIF management team are highly essential even if part of the context is redundant sometimes. The opportunity to meet with GBIF team is invaluable and very informative. We would suggest everyone that receive BIFA grant in the future to not skip a single meeting.

### Post Project Activity(ies)

Post project activities beside of finishing image capturing of the rest 1,500 bee specimens to update the database and attaching QR codes to every specimens in the collection are:

- (1) Applying for a BIFA6 data mobilization grant, which was already submitted under the project BIFA6\_009: Bee data mobilization from TIGER project. This proposed project is aim to extend our current Thai bee database to include specimens that were collected from prestigious and difficult to access national parks and wildlife sanctuaries in Thailand. This adds-on project will provide invaluable information on the occurrence records of bees in protected areas collected through passive collecting method, i.e., malaise trap, in Thailand, which are largely lack in our current dataset.
- (2) The current Thai bee dataset will provide a backbone information for future work in DNA barcoding of Thai bees associated with data in iBOL (International Barcoding of Life) and GBIF. We are planning to recruit funding of this project using the dataset we provided to GBIF.
- (3) Continue working with GBIF on this current database and expanding our collection and the occurrence records beyond the number earlier committed. As time progress, the number of specimens in the database will be updated and should GBIF permit, we will continue adding more records into the dataset to keep the dataset active and up to date with new information.

### Sustainability plans

---

The digitization of bee occurrence records will be disseminated and shared with other research collections, universities, and institutions in Thailand through communications to the NSTDA website via the Thai Bee Database (under development in <http://61.91.64.2:8075>) and other social media channels.

As mentioned earlier in Post Project Activities, we are expected to receive additional bee specimens from our collaboration with the Queen Sirikit Botanical Garden (QSBG: <http://www.qsbg.org/QSBGenglishweb/index.html>) in Chiang Mai. These specimens will be processed, curated, and digitized in the future should the project is being funded by GBIF through our proposal BIFA6\_009 and a support from NSTDA.

In addition, our current Thai bee dataset will provide a backbone information for a future work in DNA

barcoding of Thai bees. We are planning to establish a barcoding project using the dataset we procured for GBIF to enhance the capability and performance of the database. The project is aimed to collaborate with iBOL to ascertain the quality of the data generated and broad public access.

In the future, we are planning to expand our collection beyond the number we earlier committed. As time progress, the number of specimens in the database will be updated and should GBIF permit, we will continue adding more occurrence records into the dataset to keep the dataset active and up to date with new information.

---

**GBIF leads the Biodiversity Information Fund for Asia (BIFA), a programme funded by the Ministry of the Environment, Government of Japan. The programme provides supplementary support for activities addressing the needs of regional researchers and policymakers through mobilization and use of biodiversity data.**

