BIOGENESIS Jurnal Ilmiah Biologi

pISSN 2302-1616, eISSN 2580-2909
Vol 8, No. 2, December 2020, pp. 203-209
Available online http://journal.uin-alauddin.ac.id/index.php/biogenesis
DOI https://doi.org/10.24252/bio.v8i2.15697

The Digital Herbarium: Solutions for Data Collection and Identification of Indonesian Plant Diversity

Endro Setiawan[1,2]*, Dedy Darnaedi[1], Ismail Rachman[3], Teguh Triono[4,5], Campbell O. Webb[6]
[1]Department of Biology, Graduate School, Universitas Nasional
Jl. Harsono RM No. 1 Pasar Minggu, South Jakarta, Indonesia. 12550
[2]Gunung Palung National Park Office
Jl. Gajah Mada, Delta Pawan, Ketapang, West Kalimantan, Indonesia. 78811
[3]Herbarium Bogoriense, Research Center For Biology
Jl. LIPI, Bogor, West Java, Indonesia. 16911
[4]Forum Pohon Langka Indonesia (FPLI)
Jl. Ir. H. Juanda No. 18 Bogor, West Jawa, Indonesia. 16122
[5]Polytechnic Wilmar Business Indonesia (WBI)
Jl. Kapten Batu Sihombing, Medan Estate, North Sumatera, Indonesia. 20371
[6]Departemen of Biology, University of Alaska
3211 Providence Dr, Anchorage, Alaska. 99508. USA
*Email: endronesia@gmail.com

ABSTRACT. Indonesia is one of the world's biodiversity hotspots. It is estimated to be the home of 9.5% flowering plant species, making it the seventh country with the highest biodiversity. Plant data collection is necessary to ascertain the level of plant biodiversity, as such data help in conservation efforts and long-term management. One of the methods applied is the collection of plants, with the purpose to acquire as much data about its biological resources. The collected specimen are then gathered and processed into a herbarium to be used as an information source in managing biological resources. Unfortunately, there are some difficulties related to the making and management of a herbarium. This study aims to assess the advantages and disadvantages of photo-specimens (digital herbarium) for documenting plant biodiversity in Indonesia. The methods need steps including biological recording, specimen preparing, macro-mode capturing, and last stage identification. About 2149 plants have been gathered from Borneo, Seram, Waigeo, Flores and Sulawesi which consisted of 152 family, 512 genus, and 1,832 species, with a total of 30391 pictures of plant parts. From the experiment conducted on 672 specimens, it achieved 98.8 % accuracy on the family level and 80.1 % accuracy on the genus level, while the species level reached 78.8%. The results showed that digital herbarium can be used to conduct identification and data collection of plant biodiversity. Furthermore, this method is simple, cheap and relatively easier to conduct. The output is a catalog of plant species in specific areas, which provides better understanding about plant identification and biodiversity, enhances conservation practices, and provides better long-term protection for Indonesian plant biodiversity.

Keywords: digital herbarium; herbarium development; herbarium records; plant biodiversity; *Trigonachras acuta*

Article History: Received 04 September 2020; Received in revised form 01 October 2020; Accepted 17 November 2020; Available online 30 December 2020
How to Cite This Article: Setiawan E, Darnaedi D, Rachman I, Triono T, Webb CO. 2020. The digital herbarium: solutions for data collection and identification of Indonesian plant diversity. *Biogenesis: Jurnal Ilmiah Biologi*. vol 8(2): 203-209. doi: https://doi.org/10.24252/bio.v8i2.15697.

INTRODUCTION

Indonesia is one of the world's biodiversity hotspots. The country has 24632 kinds of flowering plants or 9.50% of the total species worldwide (Retnowati *et al.,* 2019) making it the seventh country with the highest biodiversity. The total size of Indonesia only amounts to 1.3% of earth size, but contains a high variety of biodiversity (Kusmana & Hikmat, 2015). The country's geographical position is very strategic, and plays a significant role.

The knowledge of biodiversity is not a prerequisite for protecting it, unlike boundaries, policies and policing, but information about biodiversity is important (Webb *et al*., 2010). One of the things which can be done to support this knowledge is species data collection, and plays an important role in identifying the level of biodiversity in Indonesia. The data collection for plants is applied in the field through plant collection. This collection of herbarium specimens constitute the main part of the activities of biology experts in the past and as well as the present (Flannery, 2013). Herbarium specimens have become a very important comparison material in the study of biodiversity.

Alternatively, there are some weaknesses of herbarium specimens which include: (a) The making process is difficult and can only be performed by experts, (b) they require specific room temperature for storage, (c) they require regular maintenance, (d) changes in plant color, (e) absence of many important characteristics, (f) limited numbers (g) inability to be shared with many people, (h) and easy damage of specimen due to inadequate maintenance or high usage for identification and manual data checking. To make it accessible to many people and protect it from damage due to high usage, the digitizing of herbarium specimens, as performed in the United States, is applicable (Barkworth & Murrell, 2012; Thiers *et al.*, 2016). Although there are still some weaknesses that need to be addressed in order to support cheap and easy collection of plant data by everyone with limited or without botanical or taxonomic background. The alternative to the physical specimens involves taking many pictures of plants with high quality in the field. These pictures are then arranged into a catalogue known as a digital herbarium. This alternative has the following advantages: a) easy making process, b) no changes in plant colors, c) easy maintenance and storage, d) important characteristics are not easily lost and e) can be shared with many people.

This study aims to assess the advantages and disadvantages of photo-specimens (digital herbarium) for documenting plant biodiversity in Indonesia. The reports were made using photographs taken from two projects: the first a multi-island, plant biodiversity project (Project XMalesia) from 2011 to 2015, and the second a citizen scientist project carried out by amateur botanists in Kalimantan Barat in 2014. Digital herbarium becomes a solution to address the limitation in making conventional herbarium, collecting the data, and identification of plant biodiversity in Indonesia. This method can be a supplement for the traditional herbarium. The existence of this method will make it easier for data collection and identification of plant biodiversity in Indonesia.

MATERIALS AND METHODS

The materials and equipment required for making the digital herbarium included the plant specimen (a branch of leaf) with a flower or fruit, camera, cutting scissors, cutter blade, black cloth as photo background, and measurement ruler for scale. In some cases, a needle or special handler was also required to create a more detailed picture. There are six steps in making the digital herbarium.

Recording the information of the plant that will be collected. A biological record contains four main pieces of information: What, Where, When and Who of what was recorded (the ‘metadata’). Where is the spatial location; When is the date or, occasionally, the range of dates over which the record was collected; Who is the person that made the record (Isaac & Pocock, 2015). The information of a specific plant is very important in data collection. Therefore, this activity involved data collection on its habitus (tree, shrubs, liana, palm, etc.), the habitat where the plant is found, the coordinate, tree identifier (bark, buttress root, sap, branches) and the number of plants, whether many or few in the finding location. Furthermore, a preliminary type of identification is also required. When collecting a photo-specimen, a picture of a cut on the tree (nick/slash the trunk) is necessary to record any special characteristics such as smell or sap.

Prepare the specimen that will have its picture taken. The specimen required for the picture is a leaf branch with complete parts. This signifies that it must also have a flower or fruit, as these are important characters in identifying plants, especially at species level (Wäldchen *et al*., 2018). The perfect specimen, for example without any part eaten by worms, is chosen, and must be complete, in order to record all its essential parts, especially the leaf veins. Leaf veins are an important characteristic for species level identification.

Taking the picture. For a camera without manual setting, pictures are usually captured in the macro mode, while cameras with manual setting require lens opening, shutter speed, white balance, lens focus and the flash strength. The setting of these parameters is necessary to achieve a perfect picture. Due to the development of smartphone technology, some smartphones could also be used to take the picture with very good quality.

The first part to have the picture taken is a branch with a good view of its leaves. For size comparison or scaling, place the measurement ruler below. Afterwards, the picture of the leaf (upper and lower part), petiole, stipules, axille, bud, flower (in various positions), flower split, fruit (from various angles) and the fruit split (horizontal or vertical) are taken. The pictures have to be taken carefully to ensure a perfect result. A good picture is usually able to show minute characteristics such as the small hairs on the leaf.

Editing the picture. The pictures that have been taken are then cropped according to the needs. It is recommended that cropping is done with appropriate scaling to ensure a final picture of the same size. Furthermore, cropping, besides removing the unnecessary part of the picture, is also useful in reducing the file size of the picture. To mark (watermark) or specify that the pictures belong to someone, each picture is renamed with special codes that can be used by everyone. Other than as a marking, the code usage helps significantly in database management, and can be called by using a query system when formulating a database. After renaming the pictures, the next process is resizing, to make it easier to upload. The original files with bigger size should still be kept, as the resized pictures usually have a lower quality. In this process, attention is required for the maximum resize limit to keep pictures sharp.

Upload. The next step is to upload, which involves the submission of plant data into the database. During the upload it is expected that the plant data is complete to make it easier for verification. After uploading, it automatically becomes part of the database.

Last stage identification. Data that has been submitted into the database is then corrected and verified. The aforementioned correction can be performed through identification or by a taxonomy expert. Factually, cooperation with taxonomy experts is needed in this identification process. After verification, the database is then published.

Fig. 1. Digital herbarium of *Trigonachras acuta* Radlk

RESULTS AND DISCUSSION

During a project to document plant biodiversity across Indonesia, 2149 physical collections of plants were gathered from Borneo, Seram, Waigeo, Flores and Sulawesi which consisted of 152 family, 512 genus, and 1832 species. Of these 2149 specimens, 30391 pictures of plant parts were obtained. Pictures of a branch of leaves, upper and lower leaf, petiole, stipules, flower, fruit and related data such as the type, name, collector, location, coordinate, date, the photographer and the identification (Fig. 1).

In a second project working with citizen scientists in West Kalimantan, 1012 more collections were made from some areas in West Borneo. These were used to assess the potential to identify and match photo-specimens to species.

The identification was conducted by Ismail Rachman of the Herbarium Bogoriense, an expert with many years of experience in identifying plants in Indonesia. From a total of 1012 specimens, the identification of a subset of 672 specimens were conducted. It was successful for 98.8% of specimens at a family level, 80.1% at genus level, and 78.8% at species level. There are some specimens that could not be identified due to sterile pictures, and some herbaceous plants, which are very difficult to identify based on only pictures. These results showed that a digital herbarium can be used to collect data and identify plant biodiversity in Indonesia. There were 109 families that were identified from 672 specimens, and the identification results from 21 families with the most genera are shown in Table 1.

Table 1. Identification test results on a digital herbarium from 21 families with the most genera.

No	Family	total_obs	only_fam_obs	only_genus_obs	to_species_obs	%_obs_to_species	species_amount
1	Phyllanthaceae	50		12	38	0.760	26
2	Rubiaceae	49	1	14	34	0.694	23
3	Dipterocarpaceae	42		7	35	0.833	25
4	Moraceae	35		6	29	0.829	24
5	Myrtaceae	32		12	20	0.625	15
6	Annonaceae	30		11	19	0.633	13
7	Euphorbiaceae	27		6	21	0.778	16
8	Meliaceae	23	1	2	20	0.870	14
9	Leguminosae	22		2	20	0.909	16
10	Clusiaceae	20		2	18	0.900	13
11	Lamiaceae	19		4	15	0.789	14
12	Melastomataceae	16		3	13	0.812	10
13	Malvaceae	16			16	1.000	10
14	Lauraceae	16		2	14	0.875	12
15	Sapindaceae	15			15	1.000	11
16	Myristicaceae	13		2	11	0.846	9
17	Apocynaceae	13	2	2	9	0.692	6
18	Elaeocarpaceae	11		5	6	0.545	4
19	Burceraceae	11			11	1.000	6
20	Primuliaceae	10			10	1.000	8
21	Anacardiaceae	10			10	1.000	7

Lesson learned from data collection by citizen scientists. The history of plant studies in Indonesia dates back to the Netherlands colonial period. Although until now Borneo which has the biggest area, is still below Java in terms of specimen numbers (Retnowati *et al.*, 2019). The lack of study and researchers of plants in areas outside of Java is one of the reasons for its low plant biodiversity. Up until 2017 there was an addition of plant types from 2014 (Widjaja *et al.*, 2014) where the addition is in the spermatophyte group totaling 5400 types, which consisted of 5.385 Angiosperms and 15 Gymnosperms (Retnowati *et al.*, 2019).

Plant data collection is also required to race against the loss of plant biodiversity. There are some factors which threaten the sustainability of plant biodiversity in Indonesia. These threats originate from the intrinsic biological factor of plants (amounts to 83%), damaged habitat (82%), over exploitation (62%) and natural disasters (6%) (Budiharta *et al*., 2011). The intrinsic biological factor of plants including small population size, limited spreading size, specific habitat, reproduction problem the need of symbiosis (Rugayah *et al*., 2017), as well as human activities, influence the arrival of new species (Trimanto & Shofiah, 2018). Furthermore, Indonesia is one of the countries with the highest levels of primary forest loss in tropical areas (Margono *et al*., 2014) which constitute the source of plant biodiversity information. The opinions of experts and modeling shows that around 15% of flowery plant species have been lost in the last few decades (Pimm *et al*., 2010; Joppa *et al*., 2011). Almost all the new species found, have also become rare locally and limited geographically (Joppa *et al*., 2011). In addition, alongside the quick loss of plant biodiversity is the occurrence of biodiversity knowledge crisis (Webb *et al.*, 2010).

This limited data collection of plant biodiversity needs to be addressed, and the digital herbarium method is one of the solutions for data collection of plant biodiversity in Indonesia. The data collection can be conducted by citizen scientists in all parts of the country. For this to work they need to be trained beforehand to ensure easy and uniform collection. Citizen scientist can be defined as a collaboration between scientist and non-scientists to collect authentic, shared and analyzed data (Jordan *et al*., 2012; Jennett *et al*., 2016). The observation conducted by citizen scientists such as plant phenology observation (Fuccillo *et al*., 2015), and invasive species (Gallo & Waitt, 2011), showed that citizen scientists with limited training can provide reliable observation when following standard and explicit protocol. In Indonesia, the role of biodiversity data collection has been conducted, such as the one by Burungnesia on bird biodiversity. They collect, analyze, archive and share data on biodiversity.

During the experiment of the digital herbarium method in West Borneo, the citizen scientists played a significant role. The experiment involved around 15 individuals with various backgrounds, jobs and education. They participated in data collection and picture taking training for the digital herbarium. After 15 days, 1020 collections were received from the field. The data collected showed that plant identification was conducted with pictures of 672 collection numbers and it achieved a high accuracy of 98.8% at family level, 80.1% at genus level, and 79.8% at the species level. This identification was conducted by an expert in this field. Therefore, the results showed that the pictures taken by citizen scientists are suitable for use in identification. The high number of plants collected also showed that the role of citizen scientists is very important for biodiversity data collection and worthy of scientific partnership. Minimally, it also serves as preliminary information data about a certain type of plant to support data collection in Indonesia, especially in areas with lack of data. Furthermore, finding an unknown species in a hotspot area will help in underlining the extraordinary biodiversity and uniqueness (Scheffers *et al.*, 2012).

Lesson learned from iNaturalist. Since 2008, iNaturalist has conducted identification for the biodiversity observations of citizen scientists. This website enables naturalists to map and share photographic observations of biodiversity around the world. Each observation consists of date, location, picture, and label containing the name of the species in the picture (Van Horn *et al*., 2017). iNaturalist possesses more than 25 million records of wildlife biodiversity with picture or audio proof, from every country, representing more than 230,000 species, collected by over 700000 individuals, alongside over 90000 individuals helping the others in identification (Seltzer *et al.,* 2019).

From the aforementioned webpage (www.iNaturalist.org) it can be seen that there are 152277 observations, with 12684 species, 3705 identifications, and 4776 observers in

Indonesia (in July 2020), which is a fantastic number. If only 15 individuals can conduct high quality data collection of 1020 plant species in a few months, this shows how much more plant data collection can be done with more individuals.

However, on further observation the pictures of plants from each iNaturalist observer have different picturing methods between one another. The standardized digital herbarium method as described above can be used to improve the situation. When the "digital herbarium method" is combined with the number of users and other benefits from iNaturalist (social aspect: following, research level confirmation, etc.) this will result in a better way to document plant biodiversity. A digital herbarium enhances the picture formats, thereby making it easier for identification. From what can be seen in iNaturalist webpage, the citizen scientists in Indonesia have high potential as an extraordinary resource for biodiversity data collection. Furthermore, picture taking is enhanced with the existence of smartphone technology which supports the capability to take pictures with high quality. Therefore, anyone may potentially document plant biodiversity from every corner of Indonesia. What is needed is better training and sharing of the 'digital herbarium method'.

CONCLUSION

The digital herbarium has the potential to solve the limitations of the traditional herbarium. The collection of plant pictures, with every step of reproduction (leaf, flower, and fruit) and its species characteristics, is an effective method as a learning tool for study and raising awareness on plant biodiversity in Indonesia. Digital herbarium in form of plant species catalogue is easier to create for specific areas, and enhances better identification and data collection of plant biodiversity in Indonesia, with an accuracy of 99.4 % at family level, 98.02 % at genus level and 76.78% at species level. The identification can be done by experts in this field, as well as through the multi-user verification method like in iNaturalist which points to ID "Research Grade". To achieve accurate results, fertile specimens (having flower or fruit) are required, as when using sterile specimens (without flower and fruit) both need to be compared with each other. This method was developed to help in plant data collection, and not to replace specimen herbarium. This is an effective method which can be used as a learning tool to develop the knowledge of plant biodiversity in Indonesia. It is simple, cheap, and relatively easy to conduct and can be used by anyone in an effort to support data collection. The output is a plant species catalogue in specific areas, which provides better identification and understanding, and enhances conservation practices to provide better long-term protection of plant biodiversity.

REFERENCES

Barkworth ME, Murrell ZE. 2012. The US virtual herbarium: Working with individual herbaria to build a national resource. *ZooKeys*. vol 209: 55–73. doi: https://doi.org/10.3897/zookeys.209.3205.

Budiharta S, Widyatmoko D, Irawati I, Wiriadinata H, Rugayah R, Partomihardjo T, Ismail I, Uji T, Keim AP, Wilson KA. 2011. The processes that threaten Indonesian plants. *Oryx*. vol 45(2): 172-179. doi: https://doi.org/10.1017/S0030605310001092.

Flannery MC. 2013. Plant collections online: Using digital herbaria in biology teaching. *Bioscene: Journal of College Biology Teaching*. vol 39(1): 1–7.

Gallo T, Waitt D. 2011. Creating a successful citizen science model to detect and report invasive species. *BioScience*. vol 61(6):459-465. doi: https://doi.org/10.1525/bio.2011.61.6.8.

Isaac NJB, Pocock MJO. 2015. Bias and information in biological records. *Biological Journal of the Linnean Society*. vol 115(3): 522–531. doi: https://doi.org/10.1111/bij.12532.

Joppa LN, Roberts DL, Pimm SL. 2011. How many species of flowering plants are there? In: *Proceedings of the Royal Society B: Biological Sciences*. vol 278(1705): 554-559. doi: https://doi.org/10.1098/rspb.2010.1004.

Jennett C, Kloetzer L, Schneider D. 2016. Motivations, learning and creativity in online citizen science. *Journal of Science Communication*. vol 15(3): 1–23. doi: https://doi.org/10.22323/2.15030205.

Jordan RC, Ballard HL, Phillips TB. 2012. Key issues and new approaches for evaluating citizen-science learning outcomes. *Frontiers in Ecology and the Environment*. vol 10(6): 307–309. doi: https://doi.org/10.1890/110280.

Fuccillo KK, Crimmins TM, de Rivera CE, Elder TS. 2015. Assessing accuracy in citizen science-based plant phenology monitoring. *International Journal of Biometeorology*. vol 59: 917–926. doi:

https://doi.org/10.1007/s00484-014-0892-7.
Kusmana C, Hikmat A. 2015. Keanekaragaman hayati flora di Indonesia. *Jurnal Pengelolaan Sumberdaya Alam Dan Lingkungan*. vol 5(2): 187–198. doi: https://doi.org/10.19081/jpsl.5.2.187.
Margono BA, Potapov P V., Turubanova S, Stolle F, Hansen MC. 2014. Primary forest cover loss in Indonesia over 2000-2012. *Nature Climate Change*. vol 4: 730–735. doi: https://doi.org/10.1038/nclimate2277.
Retnowati A, Rugayah R, Rahajoe JS, Arifiani D. 2019. Status keanekaragaman hayati Indonesia: Kekayaan jenis tumbuhan dan jamur Indonesia. Jakarta: LIPI press. pp 1–135.
Webb CO, Slik JWF, Triono T. 2010. Biodiversity inventory and informatics in Southeast Asia. *Biodiversity and Conservation*. vol 19(4): 955–972. doi: https://doi.org/10.1007/s10531-010-9817-x.
Rugayah R, Yulita KS, Arifiani D, Rustiami H, Girmansyah D. 2017. Tumbuhan langka Indonesia: 50 jenis tumbuhan terancam punah. Jakarta: LIPI press. pp 1–157.
Pimm SL, Jenkins CN, Joppa LN, Roberts DL, Russell GJ. 2010. How many endangered species remain to be discovered in Brazil?. *Natureza a Conservacao*. vol 8(1): 1–7. doi: https://doi.org/10.4322/natcon.00801011.
Scheffers BR, Joppa LN, Pimm SL, Laurance WF. 2012. What we know and don’t know about Earth’s missing biodiversity. *Trends in Ecology and Evolution*. vol 27(9): 501–510. doi: https://doi.org/10.1016/j.tree.2012.05.008.
Seltzer C. 2019. Making biodiversity data social, shareable, and scalable: Reflections on iNaturalist & citizen science. *Biodiversity Information Science and Standards*. vol 3: 1–2. doi: https://doi.org/10.3897/biss.3.46670.
Thiers BM, Tulig MC, Watson KA. 2016. Digitization of the New York Botanical Garden Herbarium. *Brittonia*. vol 68: 324–333. doi: https://doi.org/10.1007/s12228-016-9423-7.
Trimanto T, Shofiah F. 2018. Exploration of flora diversity and recommending species for reclamation of coal mining with biodiversity concept in Besiq Bermai Forest, East Borneo. *Journal of Tropical Life Science*. vol 8(2): 97–107. doi: https://doi.org/10.11594/jtls.08.02.02.
Van Horn, G., Mac Aodha O, Song Y, Cui Y, Sun C, Shepard A, Adam H, Perona P, Belongie S. 2017. The iNaturalist Species Classification and Detection Dataset-Supplementary Material. *Reptilia.* vol 32(400): 1–3.
Wäldchen J, Rzanny M, Seeland M, Mäder P. 2018. Automated plant species identification—Trends and future directions. *PLoS Computational Biology*. vol 14(4): 1–19. doi: https://doi.org/10.1371/journal.pcbi.1005993.
Widjaja EA, Rahayuningsih Y, Rahajoe JS, Ubaidillah R, Maryanto I, Walujo EB, Semiadi G. 2014. Kekinian keanekaragaman hayati Indonesia. Jakarta: LIPI Press. pp 1–344.