APPLICATION

Methods in Ecology and Evolution

The Flora Incognita app – interactive plant species identification

Patrick Mäder[1] | David Boho[1] | Michael Rzanny[2] | Marco Seeland[1] | Hans Christian Wittich[1] | Alice Deggelmann[2] | Jana Wäldchen[2]

[1]Computer and Systems Engineering Institute, Technische Universität Ilmenau, 98693 Ilmenau, Germany

[2]Max Planck Institute for Biogeochemistry, 07745 Jena, Germany

Correspondence
Patrick Mäder, Computer and Systems Engineering Institute, Technische Universität Ilmenau, 98693 Ilmenau, Germany
Email: patrick.maeder@tu-ilmenau.de

Funding information
German Federal Ministry for the Environment, Nature Conservation, Building and Nuclear Safety (BMUB) grants: 3514685C19, 3519685A08 and 3519685B08; the German Ministry of Education and Research (BMBF) grants: 01LC1319, 6PGF0334; and the Thuringian Ministry for Environment, Energy and Nature Conservation grant: 0901-44-8652; Stiftung Naturschutz Thüringen (SNT) grant: SNT-082-248-03/2014.

1 Being able to identify plant species is an important factor for understanding biodiversity and its change due to natural and anthropogenic drivers.
2 We discuss the freely available Flora Incognita app for Android, iOS, and Harmony OS devices that allows users to interactively identify plant species and capture their observations. Specifically developed deep learning algorithms, trained on an extensive repository of plant observations, classify plant images with yet unprecedented accuracy. By using this technology in a context-adaptive and interactive identification process, users are now able to reliably identify plants regardless of their botanical knowledge level.
3 Users benefit from an intuitive interface and supplementary educational materials. Captured observations in combination with their metadata provide a rich resource for researching, monitoring, and understanding plant diversity.
4 Mobile applications such as Flora Incognita stimulate the successful interplay of citizen science, conservation, and education.

KEYWORDS
automated plant species identification, citizen science, biodiversity monitoring, deep learning, environmental educational, machine learning, mobile app

GERMAN ABSTRACT

1 Die Fähigkeit, Pflanzenarten bestimmen zu können, ist eine wichtige Voraussetzung für den Schutz der biologischen Vielfalt.

2 Wir stellen die freiverfügbare Flora Incognita App vor, mit der Pflanzen interaktiv bestimmt und einer digitalen Sammlung hinzugefügt werden können. Neueste, für diese Anwendung angepasste Deep-Learning-Algorithmen, welche auf einem umfangreichen Repository von Pflanzenbeobachtungen trainiert wurden, klassifizieren Pflanzenbilder mit bisher unerreichter Genauigkeit. Durch die Nutzung dieser Technologien in einem kontextadaptiven und interaktiven Identifikationsprozess können Nutzer*innen nun unabhängig von ihrem botanischen Wissensstand zuverlässig Pflanzen identifizieren.

3 Nutzer*innen profitieren von einer intuitiven Benutzeroberfläche und ergänzenden Lernmaterialien. Erfasste Beobachtungen in Kombination mit ihren Metadaten stellen eine reichhaltige Ressource zur Erforschung, Überwachung und zum Verständnis der Pflanzenvielfalt dar.

4 Die Flora Incognita App demonstriert das erfolgreiche Zusammenspiel von Bürgerwissenschaft, Naturschutz und Bildung.

Keywords: automatisierte Pflanzenbestimmung, Bürgerwissenschaft, Biodiversitäts-Monitoring, Deep Learning, Umweltbildung, maschinelles Lernen, mobile App, iOS, Android

1 | INTRODUCTION

The global loss of biodiversity is among the most urgent environmental problems of our time, threatening to compromise stability and functioning of ecosystems (Barnosky et al., 2012; Dirzo et al., 2014; Ceballos et al., 2015). Ongoing conservation efforts require an accurate understanding of spatiotemporal patterns of biodiversity and their change over time (Chapman and Busby, 1994). Deficiencies in both the quality and the availability of biodiversity data currently prevent data-driven conservation decisions, such as land-use planning and species conservation assessments (Boakes et al., 2010; Geijzendorffer et al., 2016; Proença et al., 2017). Governments and scientific agencies typically lack the resources to fund long-term biodiversity assessments by professional scientists and therefore recruit volunteers, both beginners and experts, to meet their assessment goals (Miller-Rushing et al., 2012). Biodiversity monitoring is a labour-intensive task, heavily relying on individual expertise to correctly identify species in the field. In-situ species identification is almost impossible for untrained people and challenging even for professionals, putting it beyond the reach of many nature enthusiasts (Bonnet et al., 2018). The situation is further aggravated by the increasing shortage of skilled taxonomists (Hopkins and Freckleton, 2002; Frobel and Schlumprecht, 2014). For these reasons, there has long been interest in developing automated species identification systems (Gaston and O'Neill, 2004; Wäldchen and Mäder, 2018b). Initial image-based approaches were proposed 15 years ago but only now have become a reliable alternative to manual identifications that can reduce uncertainty and labor of identifications (Affouard et al., 2017; Wäldchen and Mäder, 2018a; Wäldchen et al., 2018; Seeland et al., 2019; Jones, 2020). Recent boosts in data availability accompanied by substantial progress in machine learning algorithms, notably deep convolutional neural networks (CNNs) (LeCun et al., 2015), pushed these approaches to a "production-ready" state. Automated species identification can now significantly contribute to biodiversity and conservation research (Bonnet et al., 2020). In this paper, we describe Flora Incognita, a mobile application for automated plant species identification and observation recording. The application combines novel advances in machine learning-enabled identification with a plant species field guide. Flora Incognita differs from alternative plant identification systems by taking a multi-modal and interactive approach that not only analyzes a single image depicting an unknown plant but incorporates habitat information and queries the user for images of one or more complementary plant organs to deliver a precise identification (Wittich et al., 2018; Rzanny et al., 2019; Seeland and Mäder, 2021).

2 | METHODS

2.1 | Interactive identification process

The first user action in the identification process is choosing which growth form the unknown species belongs to. We distinguish between forbs, grasses, ferns, and trees to request specific image perspectives e.g., "Take an image of the tree's stem." We found that by requesting these detailed perspectives, users typically take better pictures. In addition to automatic image classification, we also incorporate environmental variables in the identification process. In parallel to querying a user about the growth form of the unknown plant, we automatically transfer the occurrence's geolocation (given the user's consent) and date to predict a prior for the later image analysis. This prediction incorporates the botanical, geographical and climatic context of an observation and provides a first hypothesis about which species are more or less likely to be identified in the current observation (Wittich et al., 2018). In an adaptive number of succeeding interactions, we ask the user to take images of growth form-specific organs, such as flower or leaf. These images pass an immediate plausibility check, notifying the user if the image is unlikely to depict part of an actual plant, before being uploaded for classification. The process terminates when either a species has been identified with a sufficiently high score or no more additional information has been gained in the preceding interactions. We run intensive studies to determine which perspectives are the most characteristic per growth form and are easy to acquire for a user (Rzanny et al., 2017, 2019).

2.2 | Deep neural network classifier

The automatic identification used in Flora Incognita is based on latest machine learning technologies. We design and continuously retrain three types of models based on our repository of species observation data once a significant amount of new observations is available. First, a convolutional neural network classifier analyzes single images and predicts a ranked list of candidate species depicted on the image. The model uses an architecture with 88.9 million learnable parameters which was itself optimized using machine learning (Zoph et al., 2018) and is trained on currently more than one million plant images on a cluster of GP-GPUs (general-purpose graphics processing units) over a period of several months (full training). Second, a deep feedforward network uses location embeddings and similarity learning for predicting likely species at a given location and time based on presence-absence maps, occurrence records acquired by our users, and various databases, e.g. soil type, landcover, phenological regions. Finally, we train a recurrent neural network model with structured observations to learn an optimal fusion of the different information sources to predict a candidate species. These observations consist of multiple feature vectors extracted from the images depicting different plant organs and different perspectives of the individual by the aforementioned networks.

2.3 | The Flora Incognita taxonomy

We initially focussed on native wild growing plants in Germany and chose a widely accepted local list of ferns and vascular plants as a taxonomic backbone (Wisskirchen and Haeupler, 1998). We adopted all included taxa at species rank except for the genera *Taraxacum*, *Sorbus*, *Rubus*, *Pilosella*, *Hieracium*, and *Oenothera*. Species within these genera are

FIGURE 1 Conceptional overview of the Flora Incognita ecosystem. Three distinct data flows integrate the system's components. The **identification flow** handles user identification requests and records identified plants as observations. The **validation flow** continuously involves a team of expert botanists to review rare and critical observations, preparing new data for the next training cycle. Upon sufficient new validated observations, a **training flow** retrains the identification service aiming for improved accuracy.

challenging even for experts to identify with certainty, and therefore it is exceptionally difficult to acquire trustworthy training data required to develop accurate deeper within-app resolution. When evolving our taxonomy towards a more holistic flora covering species around the world, including those occurring in parks and gardens, we migrated to the Cataloque of Life (CoL) while maintaining the non-resolved genera as discussed above. Relying on CoL allows us to easily include new species and species groups. Furthermore, it simplifies a future data exchange with biodiversity platforms like the Global Biodiversity Information Facility (GBIF) that also derive their taxonomy from CoL. In the future, we aim to achieve greater taxonomic resolution of the non-resolved genera by engaging experts to contribute trustworthy observations of the respective species.

3 | RESULTS

3.1 | The Flora Incognita ecosystem

We designed the Flora Incognita system as a flexible client-server solution consisting of scalable micro-services running in our data center and client applications making the identification service accessible in different usage scenarios. Conceptually, the server side consists of an observation service, an identification service, and a training service (cp. Fig. 1). The observation service handles user-generated observation records by storing them in a repository, making them available across a user's different devices, and providing them for retraining of the identification system. The identification service realizes interactive species identification as used by our Flora Incognita app as well as batch-wise identification for complete observations already containing all data. The training services continuously retrains our identification system once a significant amount of new and manually reviewed observation data is available. Our multi-platform client software ecosystem currently consists of three apps freely available for Android, iOS, and Harmony OS and is developed using open-source web application frameworks which enable us to maintain a modular codebase that ensures maximum reuse and consistency of functionality across applications. Our Flora Incognita app provides an interactive process that adaptively guides a user to a desired identification. In contrast, our Flora Capture app (Boho et al., 2020) encourages users to take multi-image observations which are batch-wise identified upon sync to our server, providing a digital herbarium. Finally, our Flora Expert app allows users to review observations and is currently only available to invited botanists. Additionally, we provide an application programming interface (API) for registered external clients that allows other apps and services to incorporate our species identification.

3.2 | Application details

The Flora Incognita system currently allows users to automatically identify 4,851 vascular plant species. The app was launched in April 2018 and is freely available in 19 languages for Android, iOS, and Harmony OS devices. It has been installed more than 2.75 million times around the globe. The Flora Incognita app consists of four main usage scenarios accessible from the app's home screen (cp. Fig. 2):

FIGURE 2 Screenshots from the Flora Incognita App, including a) the species list with corresponding fact sheet, b) the identification workflow, c) the news page and d) the observation list.

Identify plant: The app guides users through an adaptive process of taking one or more images depicting specific organs of an unknown plant, such as a flower or a leaf. Which images are requested and in which order is determined automatically based on an observation's context, i.e., the growth form of the unknown plant, the current season and already acquired information. The identification process requires an internet connection for transferring images and metadata to the server and receiving results. In areas without network coverage, users can take images of an unknown plant with the device's camera and later import them into the identification process, in addition to all important metadata pertaining to that image (i.e., date and location). All images are analyzed using a cascade of deep neural networks (cp. Figure 1) on the Flora Incognita computer cluster. The app will either suggest a single plant species or a short list of similar species ranked by identification probability. For each species, a comprehensive fact sheet and informative images depicting different perspectives and organs are provided. Users are requested to confirm the correct species at the end of the process to commit an observation.

My observations: After the user has confirmed the observation during the identification process, the observation is stored in the personal plant list. This plant list can be searched, sorted, and exported in comma-separated values format (CSV). Upon selecting an individual observation, users can review all details, revise the identified taxon and share the observation via social media.

Species list: The species list shows all species currently identifiable with the system and can be searched by scientific name, common name, genus and life form or a combination thereof. For each species, a comprehensive fact sheet provides information about its characteristics, ecology, toxicity, and status. These fact sheets are automatically updated to the user's current location meaning that they include ecological, protection and distribution information relevant at this position. Additionally, the fact sheets contain links to national floristic websites providing more in-depth information that goes beyond the average user interest. While we aim for high-quality fact sheets across all supported languages and for a global geographic range, not all information is available yet. We are continuously adding content and have designed a solution that facilitates complete flexibility. For example, we are collaborating with KOSMOS, a publisher of widely renowned analog field guides, to integrate their species fact sheets for customers who purchased field guides from the "Was blüht denn da?" series (Spohn et al., 2020). Future collaborations include the development of children-specific fact sheets. Where multiple fact sheets are available, the user can select the desired one via the app's settings.

News: In the news blog, we discuss app-related topics, e.g., how to take pictures most suitable for automated identification and keep users informed with frequently updated stories about topics related to plant diversity.

3.3 | Identification accuracy

We use a benchmark holdout dataset of non-trained images to continuously evaluate our solution. Our most recent classifier alone identifies the 4,851 supported plant species with a taxon-averaged accuracy of 83% based on single images. Furthermore, we studied the overall performance when acquiring, where necessary, multiple images and analyzing location and time of an observation. We asked two expert botanists to manually check the same 1,000 randomly drawn real user observations from which they felt able to assess 847 based on the the available images. They found that 93% (787) of the observations were correctly and 7% (60) incorrectly identified by our app. Furthermore, they found that the majority of the 7% confused observations were cultivated relatives mixed up with the wild living species

FIGURE 3 (a) Distribution of *Bunias orientalis* observations across Germany, where this species is considered invasive (source of the underlying map: Leaflet | OpenStreetMap contributors, CC-BY-SA). (b) Observations of five common plant species shown as the fraction of total species observation per day in 2019 and 2020.

supported by our app. We are continuously expanding the set of supported species to overcome these limitations. Table 1 compares Flora Incognita with two other free research-grade plant identification apps. All three are developed within a scientific context and aim to identify plant species and document their occurrences. Results of the four referenced comparative studies show that Flora Incognita achieves state of the art identification accuracy. We attribute Flora Incognita's varying performance mainly to the studies' experimental protocol, e.g., solely using single image observations (Jones, 2020; Shapovalov et al., 2020), unavailable or wrong geolocation preventing habitat analysis (Jones, 2020), and identifying non-supported taxa (Schmidt and Steinecke, 2019; Jones, 2020). However, Jones (2020) still concludes that the Flora Incognita app is a very valuable tool even for botanists and ecologists during field studies.

	# supported	Geographic	Google Play Metrics**		Comparative evaluations			
App	species	focus	installs	ratings	[1]	[2]	[3]***	[4]***
Pl@ntNet	30,261 *	western Europe + focus projects	10M+	139,968	76.7%	78.0%	55.0%	52.1/100
iNaturalist	N/A	world-wide	1M+	5,494	70.0%	–	–	60.7/100
Flora Incognita	4,851	central Europe	1M+	9,504	80.0%	87.8%	71.0%	60.3/100

TABLE 1 Comparison of research-grade plant identification apps in terms of focus, distribution, and performance, where [1]: (Schmidt and Steinecke, 2019), [2]: (Lüdemann, 2020), [3]: (Shapovalov et al., 2020), and [4]: (Jones, 2020). *(August et al., 2020). **We solely report Google Play metrics as of January 2021 since Apple's AppStore only reports ratings per local store and does not publicly show installs per app. *** Uses single image observations (captured from a computer screen (Jones, 2020)) to gain a repeatable experimental setup but thereby neglecting our multi-image and context analyses that have been demonstrated to substantially improve identification accuracy (Rzanny et al., 2019; Seeland and Mäder, 2021).

4 | DISCUSSION

Flora Incognita can help to detect important biological indicators for local environmental changes by providing a spatially and temporally referenced series of species occurrences. When developing Flora Incognita, we found that there is a great need for and interest in better technology to acquire biodiversity data by citizens, professional scientists and educational practitioners. Future sources of monitoring data will include semi-automatically and automatically captured data covering large spatial scales. Since species records are obtained in return for identifying plants that Flora Incognita's users are interested in, i.e. crowdsourced, the collected biodiversity data is opportunistic and potential biases need to be considered for analysis. This user focus is reflected in the recorded plant species, e.g., we found frequency of a species' observations to be related to user bias, with ubiquitous and conspicuous species being recorded more frequently than rare and cryptic species. Hence, the most frequently observed species represent broadly distributed, often ruderal and nitrophilic species. At the same time, more observations are collected in densely populated areas along roadsides and much less in remote poorly accessible areas away from paths and roads. Despite such peculiarities, the sheer amount of collected plant observations helps overcome these biases when data are appropriately handled in the analysis. In Mahecha et al. (2021), we show that after accounting for the most prominent biases, Flora Incognita observations from a single vegetation season are already sufficient to reconstruct well-known biogeographical patterns. This is a clear indication that collected observations are meaningful and can indeed support monitoring.

One application scenario for these data is the monitoring of invasive species, which are a major threat to biodiversity associated with high economic cost for our society (estimated at nearly 12 billion Euro per year in Europe) (Jeschke et al., 2014; Weber and Gut, 2004). Invasive species are often conspicuous, attract human interest, and can be easily identified by an automated approach. Figure 3 (a) illustrates the potential of our data for assessing spatial occurrence patterns of *Bunias orientalis*, an invasive neophyte in Germany. The species spreads rapidly and is replacing native and rare plant species from species-rich meadow and semi-dry grassland biotopes. Early detection and rapid response are critical processes to prevent the spread and establishment of such invasive species (European Union, 2014). The Flora Incognita system provides up-to-date and high-resolution occurrence data. Nature conservation authorities already use this data to quickly initiate control measures of multiple invasive species. Furthermore, Flora Incognita can be used to launch citizen science projects studying research questions focussed on biodiversity in urban and agricultural areas. The strength of such a system lies in producing long-term data records covering a large variety of species in high spatial dimension.

Another application scenario is the monitoring of phenology, an important bioindicator for climate change. Barve et al. (2020) found that plant observations recorded by citizen scientists allowed for a fine-scale delimitation of phenological events while not being specifically captured for this purpose. While being an unsystematic assessment, the large numbers of contributing users might still deliver considerable samples that can supplement systematic monitoring protocols. Future iterations of this research will identify species that are suitable for phenological monitoring conducted via citizen science, i.e., being often and unambiguously recorded and having short flowering and fruiting phases. Since Flora Incognita's observations document the exact timing of a record, it is possible to differentiate them within the growing season and to obtain phenologies for individual species. Figure 3 (b) reports observations for five common plant species as fraction of total species observations per day in 2019 and 2020 in Germany. The observation patterns correlate with the expected phenologies of the particular species. Additionally, we find that the blooming of *Galanthis nivalis*, *Fiaria verna* and *Alliaria petiolata* took place earlier in 2020 than in 2019, while we find the contrary for *Daucus carota*. Furthermore, certain species, such as *Crataegus monogyna* in the figure, show two observation peaks (i.e., flowering, fruiting). With this information, we can deduce distinct phenological phases and perform detailed spatial and temporal analyses of flowering plant biodiversity. With a continuously growing database, we expect new insights into inter-annual variability of phenology in response to climate dynamics in the future.

Furthermore, automated identification has an educational aspect. From a didactic perspective, the Flora Incognita app is already an established means to teach plant diversity in schools, universities, and environmental education institutions. Although the app automates the identification process, it can still assist users in acquiring species knowledge and answering questions about biodiversity, ecological facts, and relationships of species. Over time, it teaches users in an experimental manner the gestalt for how to identify, even if they do not have this knowledge on their own. In the future, we aim to provide student-targeted species fact sheets highlighting species' characteristic traits and helping students to improve their skills. Last but not least, we found that the proposed intuitive identification process stimulates social awareness for plant diversity and biodiversity. This observation stems from thousands of user reviews across the respective app stores and a continuous stream of feedback through our support channels. Comments such as “At last, we're not just 'blindly' walking through the forest!” or “This is a fun way of finding out more about the environment” show that the app is making an important contribution towards raising awareness of biodiversity. Awareness leads to initiatives and budgets for biodiversity conservation, which is one of the main goals of the United Nations (1992)

Convention on Biological Diversity (CBD) which emphasizes the importance of public education and awareness as a crucial tool (cp. CBD's article 13). Furthermore, providing convincing, appealing, and well implemented applications might inspire previously uninvolved citizen scientists to participate in biodiversity initiatives.

REFERENCES

Affouard, A., Goëau, H., Bonnet, P., Lombardo, J.-C. and Joly, A. (2017) Pl@ ntnet app in the era of deep learning. In *ICLR: International Conference on Learning Representations*. Toulon, France.

August, T. A., Pescott, O. L., Joly, A. and Bonnet, P. (2020) Ai naturalists might hold the key to unlocking biodiversity data in social media imagery. *Patterns*, **1**, 100116.

Barnosky, A. D., Hadly, E. A., Bascompte, J., Berlow, E. L., Brown, J. H., Fortelius, M., Getz, W. M., Harte, J., Hastings, A., Marquet, P. A. et al. (2012) Approaching a state shift in earth's biosphere. *Nature*, **486**, 52.

Barve, V. V., Brenskelle, L., Li, D., Stucky, B. J., Barve, N. V., Hantak, M. M., McLean, B. S., Paluh, D. J., Oswald, J. A., Belitz, M. W. et al. (2020) Methods for broad-scale plant phenology assessments using citizen scientists' photographs. *Applications in Plant Sciences*, **8**.

Boakes, E. H., McGowan, P. J., Fuller, R. A., Chang-qing, D., Clark, N. E., O'Connor, K. and Mace, G. M. (2010) Distorted views of biodiversity: spatial and temporal bias in species occurrence data. *PLoS Biology*, **8**, e1000385.

Boho, D., Rzanny, M., Wäldchen, J., Nitsche, F., Deggelmann, A., Wittich, H. C., Seeland, M. and Mäder, P. (2020) Flora capture: a citizen science application for collecting structured plant observations. *BMC Bioinformatics*, **21**, 1–11.

Bonnet, P., Goëau, H., Hang, S. T., Lasseck, M., Šulc, M., Malécot, V., Jauzein, P., Melet, J.-C., You, C. and Joly, A. (2018) Plant identification: experts vs. machines in the era of deep learning. In *Multimedia Tools and Applications for Environmental & Biodiversity Informatics*, 131–149. Springer.

Bonnet, P., Joly, A., Faton, J.-M., Brown, S., Kimiti, D., Deneu, B., Servajean, M., Affouard, A., Lombardo, J.-C., Mary, L. et al. (2020) How citizen scientists contribute to monitor protected areas thanks to automatic plant identification tools. *Ecological Solutions and Evidence*, **1**, e12023.

Ceballos, G., Ehrlich, P. R., Barnosky, A. D., García, A., Pringle, R. M. and Palmer, T. M. (2015) Accelerated modern human–induced species losses: Entering the sixth mass extinction. *Science Advances*, **1**, e1400253.

Chapman, A. D. and Busby, J. R. (1994) Linking plant species information to continental biodiversity inventory, climate modeling and environmental monitoring. In *Mapping the Diversity of Nature*, 179–195. Springer.

Dirzo, R., Young, H. S., Galetti, M., Ceballos, G., Isaac, N. J. and Collen, B. (2014) Defaunation in the anthropocene. *Science*, **345**, 401–406.

European Union (2014) Regulation (EU) No 1143/2014 of the European Parliament and of the Council of 22 October 2014 on the prevention and management of the introduction and spread of invasive alien species. OJ L 317 (4.11.2014), , 35–55.

Frobel, K. and Schlumprecht, H. (2014) Erosion der Artenkenner. *Abschlussbericht im Auftrag des BUND Naturschutz in Bayern eV, Nürnberg*.

Gaston, K. J. and O'Neill, M. A. (2004) Automated species identification: why not? *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, **359**, 655–667.

Geijzendorffer, I. R., Regan, E. C., Pereira, H. M., Brotons, L., Brummitt, N., Gavish, Y., Haase, P., Martin, C. S., Mihoub, J.-B., Secades, C. et al. (2016) Bridging the gap between biodiversity data and policy reporting needs: An essential biodiversity variables perspective. *Journal of Applied Ecology*, **53**, 1341–1350.

Hopkins, G. W. and Freckleton, R. P. (2002) Declines in the numbers of amateur and professional taxonomists: implications for conservation. *Animal Conservation*, **5**, 245–249.

Jeschke, J. M., Bacher, S., Blackburn, T. M., Dick, J. T., Essl, F., Evans, T., Gaertner, M., Hulme, P. E., Kühn, I., Mrugała, A. et al. (2014) Defining the impact of non-native species. *Conservation Biology*, **28**, 1188–1194.

Jones, H. G. (2020) What plant is that? Tests of automated image recognition apps for plant identification on plants from the British flora. *AoB Plants*, **12**, plaa052.

LeCun, Y., Bengio, Y. and Hinton, G. (2015) Deep learning. *Nature*, **521**, 436.

Lüdemann, J. (2020) *The usability of plant identification apps and their effect on the user's engagement with species.* Master's thesis, Leuphana University of Lünebrug.

Mahecha, M. D., Rzanny, M., Kraemer, G., Mäder, P., Seeland, M. and Wäldchen, J. (2021) Crowed sourced species occurrence data reveal macro ecological patterns. *Ecography*, **in revision**.

Miller-Rushing, A., Primack, R. and Bonney, R. (2012) The history of public participation in ecological research. *Frontiers in Ecology and the Environment*, **10**, 285–290.

Proença, V., Martin, L. J., Pereira, H. M., Fernandez, M., McRae, L., Belnap, J., Böhm, M., Brummitt, N., García-Moreno, J., Gregory, R. D. et al. (2017) Global biodiversity monitoring: from data sources to essential biodiversity variables. *Biological Conservation*, **213**, 256–263.

Rzanny, M., Mäder, P., Deggelmann, A., Chen, M. and Wäldchen, J. (2019) Flowers, leaves or both? How to obtain suitable images for automated plant identification. *Plant Methods*, **15**, 77.

Rzanny, M., Seeland, M., Wäldchen, J. and Mäder, P. (2017) Acquiring and preprocessing leaf images for automated plant identification: understanding the tradeoff between effort and information gain. *Plant Methods*, **13**, 97.

Schmidt, M. and Steinecke, H. (2019) Heimische Pflanzen mit dem Smartphone bestimmen–ein Praxistest. *Der Palmengarten*, **83**, 138–140.

Seeland, M. and Mäder, P. (2021) Multi-view classification with convolutional neural networks. *Plos one*, **16**, e0245230.

Seeland, M., Rzanny, M., Boho, D., Wäldchen, J. and Mäder, P. (2019) Image-based classification of plant genus and family for trained and untrained plant species. *BMC Bioinformatics*, **20**, 4.

Shapovalov, V. B., Megalinska, A. P., Andruszkiewicz, F. and Dołhańczuk-Śródka, A. (2020) Assessment of mobile phone applications feasibility on plant recognition: comparison with google lens AR-app.

Spohn, M., Golte-Bechtle, M. and Spohn, R. (2020) *Was blüht denn da?* Kosmos.

United Nations (1992) Convention on biological diversity. Rio de Janeiro, Brazil. United Nations.

Wäldchen, J. and Mäder, P. (2018a) Machine learning for image based species identification. *Methods in Ecology and Evolution*, **9**, 2216–2225.

— (2018b) Plant species identification using computer vision techniques: A systematic literature review. *Archives of Computational Methods in Engineering*, **25**, 507–543.

Wäldchen, J., Rzanny, M., Seeland, M. and Mäder, P. (2018) Automated plant species identification—trends and future directions. *PLoS Computational Biology*, **14**, e1005993.

Weber, E. and Gut, D. (2004) Assessing the risk of potentially invasive plant species in central europe. *Journal for Nature Conservation*, **12**, 171–179.

Wisskirchen, H. and Haeupler, H. (1998) *Standardliste der Farn- und Blütenpflanzen Deutschlands.* Ulmer.

Wittich, H. C., Seeland, M., Wäldchen, J., Rzanny, M. and Mäder, P. (2018) Recommending plant taxa for supporting on-site species identification. *BMC Bioinformatics*, **19**, 190:1–190:17.

Zoph, B., Vasudevan, V., Shlens, J. and Le, Q. V. (2018) Learning transferable architectures for scalable image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 8697–8710.

ACKNOWLEDGEMENTS

Foremost, we thank all users of our apps Flora Incognita and Flora Capture. Plant observations and user feedback help us to steadily improve Flora Incognita in multiple ways. Further, we thank all our student helpers for their help in various ways. We would also like to express our thanks to the multitude of volunteering translators helping us to make this app available in 18 languages. Finally we thank the handling editor and two anonymous reviewers for their valuable comments which greatly improved this manuscript.

DATA AVAILABILITY

The Flora Incognita app is freely available for Android, iOS, and Harmony OS in the app stores. An apk file is available upon request via support@floraincognita.com

AUTHOR'S CONTRIBUTION

Funding acquisition: PM, JW; Programming: DB, HCW, MS, PM; App content: AD, MR, JW, MS; data analysis and visualization: PM, JW, MR; writing manuscript: PM, JW; comments on manuscript: MR, HCW. All authors read and approved the final manuscript.