

Standard Paper

Can we trust iNaturalist in lichenology? Evaluating the effectiveness and reliability of artificial intelligence in lichen identification

Silvana Munzi^{1,2} , Deborah Isocrono³  and Sonia Ravera^{4,5} 

¹Centro Interuniversitário de História das Ciências e da Tecnologia, Faculdade de Ciências, University of Lisbon, Campo Grande, Lisbon 1749-016, Portugal; ²Center for Ecology, Evolution and Environmental Changes & CHANGE - Global Change and Sustainability Institute, Faculdade de Ciências, University of Lisbon, Campo Grande, Lisbon 1749-016, Portugal; ³Department of Agricultural Forest and Food Sciences, University of Turin, Grugliasco 10095, Italy; ⁴Dipartimento di Scienze e Tecnologie Biologiche Chimiche e Farmaceutiche (STeBiCeF), Palermo, Sicily, Italy and ⁵NBFC, National Biodiversity Future Center, Palermo 90133, Sicily, Italy

Abstract

iNaturalist is a widely-utilized platform for data collection and sharing among non-professional volunteers and is widely employed in citizen science. This platform's data are also used in scientific studies for a wide range of purposes, including tracking changes in species distribution, monitoring the spread of alien-invasive species, and assessing the impacts of urbanization and land-use change on biodiversity. Lichens, due to their year-round presence on trees, soil and rocks, and their diverse shapes and colours, have captured the attention of iNaturalist users, and lichen records are widely represented on the platform. However, due to the complexity of lichen identification, the use of data collected by untrained, or poorly trained volunteers in scientific investigation poses concerns among lichenologists. To address these concerns, this study assessed the reliability of lichen identification by iNaturalist users by comparing records on the platform with identifications carried out by experts (experienced lichenologists) in three cities where citizen science projects were developed. Results of this study caution against the use of unchecked data obtained from the platform in lichenology, demonstrating substantial inconsistency between results gathered by iNaturalist users and experts.

Keywords: algorithm; biodiversity; citizen science; plant blindness; species identification; taxonomy

(Accepted 11 May 2023)

Introduction

Although the first citizen science activity is traditionally considered to be the Christmas Bird Count of the Audubon Society in 1900 (Butcher *et al.* 1990), public participation in acquisition of knowledge aimed at scientific research is a process that has only relatively recently increased in popularity. Coined as a description of a laboratory activity focused on the analysis of data on birds, collected over the years by a large number of amateurs (Bonney 1996), the expression 'citizen science' is currently used as an umbrella term for all approaches that entail the involvement of citizens and researchers in collaborative research activities.

In the last 25 years, thousands of projects have been organized involving millions of participants in the collection and/or processing of data all over the world, covering a wide range of domains. According to a survey focused on European citizen science projects (Hecker *et al.* 2018), 75.7% of the projects were related to life sciences, 11% to humanities and social sciences, 7.5% to natural sciences, 5.8% to engineering disciplines.

Citizen science is becoming an increasingly well-developed and valued approach, even at an international political level

(e.g. Bio Innovation Service 2018; Turbé *et al.* 2019). For example, when considering the research and innovation funding programme Horizon Europe, citizen science is recognized as a strategic practice to strengthen European citizens' trust in research and in its achievements and to raise the level of scientific literacy.

Among the most successful tools for collecting data in citizen science initiatives for natural science, iNaturalist (www.inaturalist.org/) is one of the most popular worldwide (Aristeidou *et al.* 2021). It is an online platform built, as a joint initiative of the California Academy of Sciences and the National Geographic Society, on the concept of mapping and sharing observations of biodiversity across the globe, as well as for seeking identification help from the community of users. iNaturalist uses artificial intelligence (computer vision systems) trained on photos and identifications uploaded on the platform itself to provide automated taxon identification suggestions. That potentially allows everyone to associate a putative name to an observed organism, without specific knowledge of the taxon.

iNaturalist has become very common and widely used: today, the iNaturalist community has more than two million registered users, who have made over 120 million observations of 400 000 species. Following registration on this social platform, available for PC and mobile devices, individuals can upload observations and identify organisms. Observations are reports of the occurrence of one organism in a given space and time, which can be supported by some media (currently images or sound) tagged with metadata such as taxonomic identification, date, associated text, tags, and

Corresponding author: Deborah Isocrono; Email: deborah.isocrono@unito.it

Cite this article: Munzi S, Isocrono D and Ravera S (2023) Can we trust iNaturalist in lichenology? Evaluating the effectiveness and reliability of artificial intelligence in lichen identification. *Lichenologist* 55, 193–201. <https://doi.org/10.1017/S0024282923000403>

© The Author(s), 2023. Published by Cambridge University Press on behalf of the British Lichen Society. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.



geographic information. The community (both experts and non-experts) can additionally add identifications of the uploaded observation. When two or more observers agree on the identification, its status moves from 'Needs ID' to 'Research-grade' (RG), where RG is the highest quality level in iNaturalist for an identification at a level lower than family. Moreover, RG observations are also aggregated in the Global Biodiversity Information Facility database (GBIF; www.gbif.org) by default.

These observations, which provide large-scale, real-time data on biodiversity, can also represent valuable resources for scientific research. Citizen science platforms are becoming a popular tool for researchers allowing scientists to access a wider data collection network (Follett & Strezov 2015). For easily identifiable groups, iNaturalist can be a valuable source of information on distribution, enabling it to be used to supplement traditional survey methods (Jackson *et al.* 2015). iNaturalist data have been used in a number of scientific studies for a wide range of purposes including tracking changes in species distribution both of common and elusive species (Ricca & Cheung 2021; Rosa *et al.* 2022), monitoring the spread of alien-invasive species (Creley & Muchlinski 2017; Hiller & Haelewaters 2019), studying the impacts of urbanization and land-use change on biodiversity (Lee *et al.* 2021), and analyzing plant phenology (Barve *et al.* 2020) and its links with climate change (Iwanycki Ahlstrand *et al.* 2022). The use of iNaturalist data is also increasing in studies on rare or endangered species both for monitoring and tracking the effects of conservation efforts (Wilson *et al.* 2020) and even to study the effects of recent pandemic events (Vardi *et al.* 2021). Citizen science and iNaturalist observations have been used to detect species previously unknown to science (Winterton 2020) and to explore species occupancy and distribution in areas with no systematic research institutions and scarce funding for science (Wangyal *et al.* 2022).

While the idea of mapping biodiversity and sharing observations is to be commended and the storage of open access data on a free platform is a powerful tool both for citizen science and research, we must consider that species identification is not an easy task. Developing the skills to identify species of any taxon can require years of practice and, often, instruments and products that are not within everyone's reach (i.e. microscopes, chemicals, chromatography equipment, literature).

In 2022, about one million observations of lichens were available worldwide on iNaturalist, which is less than 1% of the total uploaded observations, meaning lichens are still under-represented on the platform.

In lichenology, most of the species require microscopic observations and chemical tests and, even for experts, images are barely sufficient to reach a correct identification at the species level. Therefore, having thousands of lichen records reported by non-lichenologists and identified via computer vision (a type of artificial intelligence) raises concerns among the scientific community (McMullin & Allen 2022). In this work, we evaluated the effectiveness and reliability of lichen identification on iNaturalist comparing the results of lichen identification performed by iNaturalist users using the platform with results obtained by expert lichenologists.

Material and Methods

Test concept

The comparison between iNaturalist users and expert lichenologists was organized in well-defined areas of comparable size (FCUL campus in Lisbon, Grugliasco University campus in

Turin, and Botanical Garden and Villa Giulia in Palermo) in order to identify the different sources of variability and error in lichen recognition. The expert assessments aimed to provide the baselines against which to compare the iNaturalist users' performances and were not intended as replicates of the same experimental design.

The lichenological information on the sites, already on the platform, indicated significant variation between them. Within a radius of 50 km from the areas, the following data were available: 212 taxa by 933 observations from the area surrounding Lisbon; 82 taxa by 438 observations from the area surrounding Turin; 11 taxa by 20 observations from the area surrounding Palermo.

iNaturalist users received only standard information on using the platform through seminars included in specific projects at each site (+Biodiversidade@CIÊNCIAS in Lisbon, DIVERSAGROVET project in Turin, and activities organized under the guidance of the Italian Botanical Society, regional sections 'Sicilia - SBISI' and 'Piemonte e Valle d'Aosta', and Lichenology Working Group in Palermo). No specific information about lichen taxonomy was provided to the iNaturalist users.

'Tank Projects' for iNaturalist users

So-called 'Tank Projects' are three main citizen science projects carried out in Portugal and Italy. +Biodiversity-@CIÊNCIAS is a project funded by the Faculty of Sciences of the University of Lisbon (FCUL) with the aim of contributing to sustainability within the FCUL campus and in its interaction with the city of Lisbon (Portugal). The project began during 2020–21, aiming at increasing citizens' awareness of the importance of biodiversity. To involve the scientific community and the citizens who live and work in the area, the online project 'BioDiversity4All' (<https://www.biodiversity4all.org>) was created on the iNaturalist platform for data recording and monitoring.

The area selected for the activity extends over 160 ha, of which c. one quarter has tree cover and includes the FCUL campus, a public garden with a small artificial lake, the hippodrome and the sports complex of the faculty with an associated green area. The initiative was extremely successful, and the online platform so far has gathered over one million observations, referring to more than 10 000 species, with at least 8000 registered users and more than 11 000 followers on Facebook.

The DIVERSAGROVET project was conceived in 2021 by the Department of Agricultural, Forest and Food Sciences of the University of Turin (Italy) with the goal of raising citizen awareness of the value of biodiversity, allowing free access to the green areas surrounding university facilities. To map and monitor the biodiversity of the Grugliasco University campus, a specific data collection project was launched on the platform iNaturalist (<https://www.inaturalist.org/projects/diversagrovet>) involving students, citizens, and academic staff who work on the campus or live nearby.

Monitoring activities were performed on a 24-ha area, that includes green areas with both wild and cultivated plants, greenhouses, agricultural areas (orchards, vineyards and fields), a livestock farm and a small hill.

The SBISI activity was included among those proposed in 2022 by the regional sections 'Sicilia - SBISI' and 'Piemonte e Valle d'Aosta' and by the Lichenology Working Group of the Italian Botanical Society. The main purpose of this activity was to become familiar with the iNaturalist platform for collecting lichenological data. The selected areas were two contiguous urban

parks: the Botanical Garden and Villa Giulia. The Botanical Garden covers an area of more than 10 ha and hosts 1692 plant species within the collections (<http://ortobotanico.unipa.it/collezioni.html>). Mediterranean, tropical and subtropical plants are represented, in large collections, as well as by numerous specimens of exotic plants. Villa Giulia, a typical Italian garden, is a public park of 7 ha. In both parks there are several stone cultural heritage artworks.

Data collection

Data collection by SM, DI and SR was carried out by expert lichenologists with many years of experience in two steps: identification of the species in the field and checking dubious specimens in the laboratory.

The field activity lasted a total of 8–10 hours for each study area. On 5–7 July 2021, SR and SM carried out a survey of three half days in the FCUL campus in Lisbon, while in Turin, DI carried out a two-day survey (19 July and 30 September 2022) on the Grugliasco campus, and in Palermo SR dedicated half a day to each of the two parks (11 May 2022 at the Botanical Garden and 27 October 2022 at Villa Giulia).

Although the surveys were not intended to document all of the lichen species in the study areas, they were able to identify the most prevalent ones in the locations where project participants detected lichen occurrence. Identification and nomenclature were mainly based on the online keys published in ITALIC (Nimis & Martellos 2022), Clauzade & Roux (1985) and Smith *et al.* (2009).

Data mining and analysis

Data belonging to the projects ‘BioDiversity4All’, ‘DIVERSAGROVET’ and SBISI were downloaded from the iNaturalist platform selecting the category ‘Fungi including lichens’ for the period prior to the surveys made by the three expert lichenologists. All the records collected before the experts surveyed the areas were checked to select those referring to lichens, looking at the images when the taxonomic rank reported in the platform was insufficient to retrieve the information. When corrections were made by other users, the original identification of the first observer was considered.

Finally, a comparison was made between the species lists produced by the experts and the iNaturalist users. Taxa which were reported by experts and were not reported by iNaturalist users were investigated to determine potential issues that prevented their identification by the iNaturalist users. For species only reported by iNaturalist users the images were analyzed to evaluate, when possible, the accuracy of their identification. When images were not sufficiently clear, the ecology and distribution of the suggested species were considered to assess the reliability of their occurrence.

Results

In the three study areas, 108 species were inventoried (Table 1). The number of species that required verification in the laboratory was 13 in Lisbon, 8 in Turin and 5 in Palermo.

The experienced lichenologists (herein known as ‘experts’) identified almost 60% of the species in Turin and Lisbon, but only 33% in Palermo, when the number of species is considered but not the accuracy of the identification (Table 2). The overlap between the experts and iNaturalist users in Palermo was similarly lower, indicating a higher level of error in the users’

Table 1. A list of lichen species recorded by experienced lichenologists (Exp), iNaturalist users (iNat) and both (Both) in Lisbon, Turin and Palermo. * = species identified after laboratory verification by experienced lichenologists.

Species	Lisbon	Turin	Palermo
<i>Acarospora fuscata</i> (Schrad.) Arnold	-	-	iNat
<i>Amandinea punctata</i> (Hoffm.) Coppins & Scheid.	Exp*	-	-
<i>Anema nummularium</i> (Durieu & Mont.) Forssell	-	-	Exp*
<i>Arthonia radiata</i> (Pers.) Ach.	Exp	Exp	-
<i>Aspicilia cinerea</i> (L.) Körb.	-	-	iNat
<i>Athallia cerinella</i> (Nyl.) Arup <i>et al.</i>	Exp*	Exp	Exp*
<i>A. pyracea</i> (Ach.) Arup <i>et al.</i>	-	Exp	-
<i>Calogaya pusilla</i> (A. Massal.) Arup <i>et al.</i>	-	-	iNat
<i>C. saxicola</i> (Hoffm.) Vondrák	-	iNat	iNat
<i>Caloplaca cerina</i> (Hedw.) Th. Fr. s. lat.	-	-	Both
<i>Candelaria concolor</i> (Dicks.) Stein	Both	Both	iNat
<i>Candelariella aurella</i> (Hoffm.) Zahlbr.	-	Both	Both
<i>C. efflorescens</i> R. C. Harris & W. R. Buck	Exp	iNat	-
<i>C. reflexa</i> (Nyl.) Lettau	-	Exp	-
<i>C. vitellina</i> (Hoffm.) Müll. Arg.	-	Both	Exp
<i>C. xanthostigma</i> (Ach.) Lettau	-	Exp	-
<i>Catillaria nigroclavata</i> (Nyl.) J. Steiner	Exp*	Exp	-
<i>Cetraria islandica</i> (L.) Ach.	-	-	iNat
<i>Chaenotheca ferruginea</i> (Sm.) Mig.	-	-	iNat
<i>Chrysothrix caesia</i> (Flot.) Ertz & Tehler	-	-	iNat
<i>C. candelaris</i> (L.) J.R. Laundon	iNat	-	-
<i>Circinaria calcarea</i> (L.) A. Nordin <i>et al.</i>	-	--	Exp
<i>C. contorta</i> (Hoffm.) A. Nordin <i>et al.</i>	-	iNat	iNat
<i>Diploicia canescens</i> (Dicks.) A. Massal.	Both	-	Both
<i>Diplotomma hedinii</i> (H. Magn.) P. Clerc & Cl. Roux	-	Exp	-
<i>Flavoparmelia caperata</i> (L.) Hale	iNat	Exp	iNat
<i>F. soredians</i> (Nyl.) Hale	Both	-	-
<i>Flavoplaca citrina</i> (Hoffm.) Arup <i>et al.</i>	-	Exp	-
<i>F. flavocitrina</i> (Nyl.) Arup <i>et al.</i>	Exp	-	-
<i>Gyalolechia flavovirescens</i> (Wulfen) Söchting <i>et al.</i>	-	-	iNat
<i>Hyperphyscia adglutinata</i> (Flörke) H. Mayrhofer & Poelt	Exp	Exp	Exp
<i>Hypocomyce scalaris</i> (Ach.) M. Choisy	-	-	iNat
<i>Lecania cyrtella</i> (Ach.) Th. Fr.	Exp*	-	-
<i>Lecanora albella</i> (Pers.) Ach.	Exp	-	-
<i>L. allophana</i> (Ach.) Nyl.	-	Exp	-
<i>L. campestris</i> (Schaer.) Hue	-	-	iNat
<i>L. chlarotera</i> Nyl.	Exp*	Exp	-
<i>L. conizaeoides</i> Cromb.	Exp*	-	-
<i>L. horiza</i> (Ach.) Linds.	Exp*	-	Exp*

(Continued)

Table 1. (Continued)

Species	Lisbon	Turin	Palermo
<i>L. hybocarpa</i> (Tuck.) Brodo	-	-	iNat
<i>L. strobilina</i> (Spreng.) Kieff.	iNat	-	iNat
<i>Lecidella elaeochroma</i> (Ach.) M. Choisy f. <i>elaeochroma</i>	Both*	Both	Exp*
<i>Lepra amara</i> (Ach.) Hafellner	-	-	iNat
<i>Lepraria finkii</i> (B. de Lesd.) R.C. Harris	-	-	iNat
<i>Leprocaulon quisquiliare</i> (Leers) M. Choisy	iNat	-	-
<i>Melanohalea elegantula</i> (Zahlbr.) O. Blanco et al.	-	Exp	-
<i>M. exasperatula</i> (Nyl.) O. Blanco et al.	-	Exp	-
<i>Micarea meridionalis</i> van den Boom et al.	Exp*	-	-
<i>Myelochroa aurulenta</i> (Tuck.) Elix & Hale	-	-	iNat
<i>Myriolecis albescens</i> (Hoffm.) Śliwa et al.	-	-	iNat
<i>M. dispersa</i> (Pers.) Śliwa et al.	-	Exp	Both
<i>M. hagenii</i> (Ach.) Śliwa et al.	-	-	Exp
<i>Naetrocymbe punctiformis</i> (Pers.) R.C. Harris	-	Exp	-
<i>Naevia punctiformis</i> (Ach.) A. Massal.	Exp*	-	-
<i>Normandina pulchella</i> (Borrer) Nyl.	Exp	-	-
<i>Ocellomma picconianum</i> (Bagl.) Ertz & Tehler	-	-	Exp*
<i>Pachnolepia pruinata</i> (Pers.) Frisch & G. Thor	Exp*	-	-
<i>Parmelia sulcata</i> Taylor	-	Both	-
<i>Parmeliella triptophylla</i> (Ach.) Müll. Arg.	-	-	iNat
<i>Parmelina tiliacea</i> (Hoffm.) Hale	-	Both	iNat
<i>Parmotrema hypoleucinum</i> (J. Steiner) Hale	Both	-	-
<i>P. perlatum</i> (Huds.) M. Choisy	Exp	-	-
<i>P. pseudoreticulatum</i> (Tav.) Hale	Exp	-	-
<i>P. reticulatum</i> (Taylor) M. Choisy	Both	-	-
<i>Pertusaria heterochroa</i> (Müll. Arg.) Erichsen	Exp*	-	-
<i>Phaeophyscia hirsuta</i> (Mereschk.) Essl.	Exp	-	iNat
<i>P. insignis</i> (Mereschk.) Moberg	Exp	-	-
<i>P. orbicularis</i> (Neck.) Moberg	Both	Both	Both
<i>Phlyctis argena</i> (Spreng.) Flot.	iNat	iNat	-
<i>Physcia adscendens</i> H. Olivier	Exp	Both	Exp
<i>P. americana</i> G. Merr.	-	-	iNat
<i>P. biziana</i> (A. Massal.) Zahlbr. var. <i>biziana</i>	-	-	Exp
<i>P. biziana</i> var. <i>leptophylla</i> Vězda	-	-	Exp
<i>P. caesia</i> (Hoffm.) Fűrnr. var. <i>caesia</i>	-	-	iNat
<i>P. clementei</i> (Turner) Lyngé	Exp	-	-
<i>P. erumpens</i> Moberg	Exp	-	-
<i>P. leptalea</i> (Ach.) DC.	Exp	-	-
<i>P. millegrana</i> Degel.	-	-	iNat
<i>P. stellaris</i> (L.) Nyl.	-	-	iNat

(Continued)

Table 1. (Continued)

Species	Lisbon	Turin	Palermo
<i>P. tribacioides</i> Nyl.	Both	-	-
<i>Physciella chloantha</i> (Ach.) Essl.	iNat	Exp	-
<i>Physconia detersa</i> (Nyl.) Poelt	-	-	iNat
<i>P. grisea</i> (Lam.) Poelt subsp. <i>grisea</i>	Both	Exp	-
<i>P. grisea</i> subsp. <i>algeriensis</i> (Flagey) Poelt	-	-	Exp
<i>Physconia perisidiosa</i> (Erichsen) Moberg	-	-	Exp
<i>Placynthium nigrum</i> (Huds.) Gray	-	-	Both
<i>Porpidia crustulata</i> (Ach.) Hertel & Knoph	-	-	iNat
<i>Protoparmeliopsis muralis</i> (Schreb.) M. Choisy var. <i>muralis</i>	-	Both	-
<i>Punctelia borleri</i> (Sm.) Krog	Both	-	-
<i>P. reddenda</i> (Stirt.) Krog	iNat	-	-
<i>P. rudecta</i> (Ach.) Krog	iNat	-	-
<i>P. subrudecta</i> (Nyl.) Krog	iNat	Exp	-
<i>Pyrrhospora querneae</i> (Dicks.) Körb.	Exp	-	-
<i>Pyxine soredata</i> (Ach.) Mont.	-	-	iNat
<i>P. subcinerea</i> Stirt.	-	-	iNat
<i>Ramalina canariensis</i> J. Steiner	Exp	-	-
<i>R. fraxinea</i> (L.) Ach.	Exp	-	-
<i>R. lacera</i> (With.) J.R. Laundon	Exp	-	-
<i>R. lusitanica</i> H. Magn.	Exp	-	-
<i>Squamarina cartilaginea</i> (With.) P. James var. <i>cartilaginea</i>	-	-	Exp
<i>Tephromela atra</i> (Huds.) Hafellner var. <i>atra</i>	-	iNat	-
<i>Variospora aurantia</i> (Pers.) Arup et al.	-	-	Exp
<i>V. flavescens</i> (Huds.) Arup et al.	iNat	-	Exp
<i>Verrucaria nigrescens</i> Pers. f. <i>nigrescens</i>	-	Exp	Exp
<i>Waynea stoechadiana</i> (Abbassi Maaf & Cl. Roux) Cl. Roux & P. Clerc	Both*	-	-
<i>Xanthomendoza fallax</i> (Hepp) Søchting et al.	-	Exp	-
<i>Xanthoria calcicola</i> Oxner	Exp	-	-
<i>X. parietina</i> (L.) Th. Fr.	Both	Both	Both

Table 2. Summary of the total number and percentage of species found by experts (i.e. experienced lichenologists), iNaturalist users (species identity not checked or confirmed by experts) and both, in the three collection sites.

	Species	Total	Expert	iNaturalist	Both
Lisbon	No.	53	31	10	12
	%		57.4	18.5	22.2
Turin	No.	36	21	5	10
	%		60.0	14.3	28.6
Palermo	No.	55	18	30	7
	%		32.7	54.5	12.7

identifications. This can be attributed to a more limited iNaturalist lichen dataset available for that area (only 20 observations were available prior to this study for Palermo and surroundings).

In all cases, more than 85% of the observations by experts led to identification at the species level and around 10% at the genus level. Conversely, less than 50% of the specimens were identified at the species level by iNaturalist users, more than 50% of identifications being referred to higher taxonomic ranks (Table 3).

Macrolichens (lichens with physical features easily seen with the naked eye) were over-represented in iNaturalist. Among the taxa which were reported only by experts, about 52%, 64% and 60% were microlichens (i.e. crustose lichens with morphological features that require microscopic observations for identification) in Lisbon, Turin, and Palermo, respectively. A sizeable percentage of overlooked species belonged to *Physciaceae* (Lisbon 23%, Turin 14% and Palermo 26%), while the rest included several genera and species (Fig. 1, Table 1).

Of the identifications at the species level exclusively reported on iNaturalist, 70% were incorrect. When images were available, while it was not always possible to accurately identify the species, it was often possible to exclude the suggested identification based on diagnostic characters. Additionally, based on the ecology and distribution of the proposed species, 7% of the identifications were considered likely to be correct, while another 7% were considered unreliable.

Discussion

iNaturalist is an easy-to-use platform that can offer an effective learning tool to citizens (Chozas *et al.* 2023). It also comes at

no cost, making it accessible to everyone. The platform's appeal is that it is accessed via mobile devices, and in particular smartphones, making it particularly appealing to the younger generation.

Our findings support the conclusions reached by McMullin & Allen (2022), cautioning against the use of unchecked data obtained from the platform in lichenology, demonstrating substantial inconsistency between results gathered by iNaturalist users and experts.

Experts dealing with tiny and taxonomically difficult organisms like fungi and insects know that species identification based on images is either impossible or at least highly uncertain (Casanovas *et al.* 2014; Prudic *et al.* 2018). Our results showed that this is as true for artificial intelligence as it is for humans. In particular, three distinct sources of uncertainty can be identified in the use of artificial intelligence for lichen identification. Firstly, the need to examine microscopic and chemical characteristics prevents the identification of entire lichen genera *in situ*. During the three surveys, experts had to complete their analysis in the laboratory, supplementing the species list and identifications established during field work. This constraint was also noted by de Groot *et al.* (2023) in their analysis of citizen science initiatives aimed at surveying arthropods in forested areas.

Several lichen species have such small and inconspicuous thalli, consisting of millimeter-sized granules/squamules or having a similar colour to the substrate, that they are almost imperceptible to the naked eye. The only indication of their occurrence may be through the presence of minute apothecia which can only be detected by experienced or highly perceptive individuals. This results in a significant underestimation of

Table 3. Summary of the total number and percentage of observations grouped by taxonomic rank found by experts (i.e. experienced lichenologists) and iNaturalist users in the three collection site.

	Observations	Lisbon		Turin		Palermo	
		Expert	iNaturalist	Expert	iNaturalist	Expert	iNaturalist
Kingdom	No.	0	3	0	10	0	3
	%	0	2.1	0	6.8	0	1.9
Phylum	No.	0	0	0	0	0	18
	%	0	0	0	0	0	11.3
Subphylum	No.	0	0	0	0	0	8
	%	0	0	0	0	0	5
Class	No.	0	12	0	37	0	14
	%	0	8.4	0	25.3	0	8.8
Order	No.	0	0	0	1	0	0
	%	0	0	0	0.7	0	0
Family	No.	1	9	0	14	0	24
	%	1.8	6.3	0	9.6	0	15.1
Sub-family	No.	1	8	0	2	0	3
	%	1.8	5.6	0	1.4	0	1.9
Genus	No.	6	43	2	10	2	14
	%	10.5	30.1	6.3	6.8	7.4	8.8
Species	No.	49	68	30	72	25	75
	%	86	47.6	93.8	49.3	92.6	47.2
Tot	No.	57	143	32	146	27	159

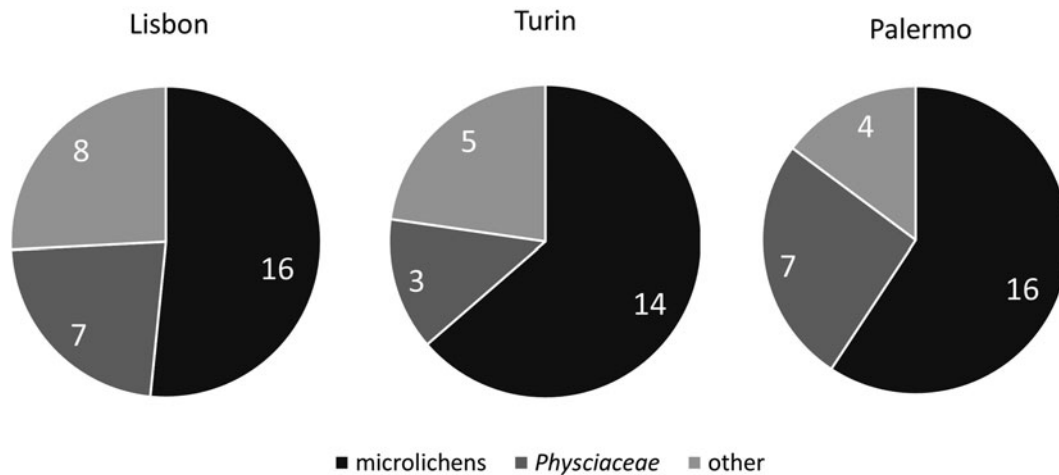


Figure 1. Proportion of overlooked lichen species belonging to *Physciaceae* in relation to groups at the three sites.

microlichens. A detection bias affecting iNaturalist observations with lichens was noted also for birds (Callaghan *et al.* 2021). Furthermore, in urban areas, where citizen science activities are more frequent but where pollution poses a significant challenge to the growth of sensitive species, poorly developed thalli may lack the morphological features necessary for accurate identification. Species that are very similar in colour, shape and size, but differ in small details or specific structures mostly ignored by iNaturalist users tend to be underrepresented, as demonstrated in this study with the *Physciaceae* family. A similar situation was reported by Koo *et al.* (2022) when non-experts were not able to discern two similar species of amphibians based on the external shape alone. Therefore, inherent characteristics of lichens present challenges in the application of iNaturalist to lichenology.

Secondly, the current functioning of the platform does not guarantee a high level of accuracy in identification. The ‘research grade’ designation is given when at least two users independently agree on the identification, regardless of their level of expertise, and there are no means of revision against it. Consequently, the accumulation of two errors in an event that can occur with high probability will result in an RG misidentification in iNaturalist records, if not properly scrutinized.

Additionally, the low quality of some images and the absence of microscopic and chemical characters make it challenging, or even impossible, for experts to provide accurate evaluations. The lack of physical specimens stored in herbaria further hinders assessment of the accuracy of the identification (McMullin & Allen 2022). Although the practice of adding images concatenated to that of the organism portrayed as a species (e.g. details, microscopic characters or images depicting their habitat and ecology) is extremely helpful in the scrutiny (Table 4), it is still an uncommon practice.

Without a strong dataset of precise identifications in a given area, as in the case of Palermo, the computer-generated suggestions may refer to species that are morphologically similar but not occurring in the area because the algorithm is not designed to take ecological factors into account when providing tentative identifications. Finally, dishonesty cannot be prevented in a system that is open access and free by definition. Those individuals who are determined to increase their personal record of observations, might blindly guess or put too much trust in the machine learning. Machine identification works well with

well-known species with large training sets, but it is unreliable for scarcely represented groups like lichens, or where background data are lacking (https://www.inaturalist.org/pages/computer_vision_demo). Even if strongly discouraged in the iNaturalist guidelines, inexperienced users tend to agree too quickly with given identifications, despite knowing almost nothing about the species.

With specific regard to the platform, some measures could help improve the possibility of posting and handling data for research purposes. A crucial point is that users should be aware of the key characters for identifying lichens. Even just taking macrolichens into consideration, there are several useful tools that could be developed, such as introductory guides, tutorials on what have to be included and documented in a useful image (e.g. important morphological features, vegetative and reproductive characters, etc.), and best practices for gathering valuable data from iNaturalist photographs. Furthermore, for the purpose of a more accurate observation, the use of a $\times 10$ magnifying lens could be encouraged. McMullin & Allen (2022) proposed a set of recommendations for documenting observations of lichens, which can also be adapted to other challenging taxa, and iNaturalist users have already taken the initiative to gather and compile online resources for this purpose. On the other hand, researchers should be facilitated in the review process of iNaturalist data. For example, creating a specific filter to separate lichens from other fungi would save a lot of time. Currently, two different phyla (*Ascomycota* and *Basidiomycota*) and various classes that do not include only lichens must be searched to obtain a dataset containing lichens. This requires labour intensive sorting and since specific filtering does not currently exist, verification must be done by checking manually all the taxa. Furthermore, considering the robustness of the data for statistical processing, redundant data should be highlighted (i.e. images of the same specimen taken by different users). In this view, also a higher agreement threshold to determine what is/is not ‘researchable’ should be reached and provided. Similarly, a higher agreement threshold for determining what is/is not ‘research grade’ could provide more reliable data. For instance, it would be useful to add ‘redundant’ identifications to records that are already RG (in fact data users can choose to limit the data set to observations with three or more confirmations, this could lead to a better population to work through).

Table 4. Assessment of the accuracy (likely, unlikely and wrong) of species identification by iNaturalist users as determined by the analysis of the associated imagery, ecology and distribution of the identified species.

	Species	Assessment	Justification
Lisbon	<i>Chrysothrix candelaris</i>	likely	Ecology: species occurring in moderately disturbed areas
	<i>Flavoparmelia caperata</i>	likely	Ecology: a mild-temperate lichen mostly found on isolated deciduous trees
	<i>Lecanora strobilina</i>	wrong	No <i>Lecanora</i> in the image. Ecology: species occurring in natural or semi-natural habitats, with no eutrophication
	<i>Leprocaulon quisquiliare</i>	wrong	Low quality image but it seems to be a <i>Lepraria</i> . Ecology: a mainly mild-temperate to Mediterranean lichen found mostly on basic siliceous rocks covered by soil
	<i>Phlyctis argena</i>	unlikely	Low quality image but the margins of the thallus appear to be different. Ecology: a subtropical to southern boreal-montane, Holarctic lichen, colonizer of smooth bark in sheltered situations
	<i>Physciella chloantha</i>	unlikely	Low quality image and thallus not well developed. Ecology: never common in heavily disturbed habitats
	<i>Punctelia reddenda</i>	likely	It appears to match the description although details cannot be confirmed
	<i>P. rudecta</i>	wrong	World distribution: America
	<i>P. subrudecta</i>	unlikely	Image of the lower surface is not available. Ecology: montane belt
	<i>Variospora flavescens</i>	likely	It seems probable; experts only considered epiphytic lichens
Turin	<i>Calogaya saxicola</i>	wrong	Thallus does not match the description. Distribution: occurrence of this species in Italy is dubious
	<i>Candelariella efflorescens</i>	likely	Most images identified as <i>C. efflorescens</i> are other species. One image is dubious. Ecology: on isolated trees, especially in orchards
	<i>Circinaria contorta</i>	wrong	All images identified as <i>C. contorta</i> are other species
	<i>Phlyctis argena</i>	wrong	All images identified as <i>P. argena</i> are other species
	<i>Tephromela atra</i>	wrong	All images identified as <i>T. atra</i> are other species
Palermo	<i>Acarospora fuscata</i>	wrong	All images identified as <i>A. fuscata</i> are other species
	<i>Aspicilia cinerea</i>	unlikely	No image present. Ecology: very rare in the Mediterranean belt in Italy
	<i>Calogaya pusilla</i>	wrong	The image identified as <i>C. pusilla</i> refers to another species
	<i>C. saxicola</i>	unlikely	The image identified as <i>C. saxicola</i> refers to another species
	<i>Candelaria concolor</i>	wrong	All images identified as <i>C. concolor</i> are other species
	<i>Candelariella aurella</i>	wrong	All images identified as <i>C. aurella</i> are other species
	<i>Cetraria islandica</i>	wrong	The image identified as <i>C. islandica</i> refers to another species
	<i>Chaenotheca ferruginea</i>	wrong	The image identified as <i>C. ferruginea</i> refers to another species
	<i>Chrysothrix caesia</i>	wrong	The image identified as <i>C. caesia</i> refers to another species
	<i>Circinaria contorta</i>	wrong	Ecology: usually it does not occur in heavily disturbed areas, mostly found at high altitude
	<i>Flavoparmelia caperata</i>	wrong	The image identified as <i>F. caperata</i> refers to another species
	<i>Gyalolechia flavovirescens</i>	likely	Low quality image. Common in the submediterranean belt
	<i>Hypocnomyce scalaris</i>	wrong	The image identified as <i>H. scalaris</i> probably refers to a moss
	<i>Lecanora campestris</i>	wrong	Low quality image but referring to another species
	<i>L. hybocarpa</i>	unlikely	Low quality image, species not recognizable without microscopic observations
	<i>L. strobilina</i>	wrong	Low quality image, species not recognizable without microscopic observations. Ecology: species occurring in natural or semi-natural habitats, with no eutrophication
	<i>Lepra amara</i>	wrong	Low quality image but referring to another species
	<i>Lepraria finkii</i>	wrong	Low quality image but referring to another species
	<i>Myelochroa aurulenta</i>	wrong	World distribution: pantropical, eastern North America and Japan

(Continued)

Table 4. (Continued)


Species	Assessment	Justification
<i>Myriolecis albescens</i>	wrong	Low quality image but referring to another species
<i>Parmeliella thriptophylla</i>	wrong	The image identified as <i>P. thriptophylla</i> refers to another species
<i>Parmelina tiliacea</i>	wrong	Not present in the image
<i>Phaeophyscia hirsuta</i>	wrong	Not present in the image; other species of <i>Physcia</i> are present
<i>Physcia caesia</i>	wrong	The image identified as <i>P. caesia</i> refers to another species
<i>P. americana</i>	wrong	World distribution: Eastern North America
<i>P. millegrana</i>	wrong	World distribution: Eastern North America
<i>P. stellaris</i>	wrong	Not present in the image, other species of <i>Physcia</i> are present
<i>Physconia detersa</i>	wrong	The image identified as <i>P. detersa</i> refers to another species
<i>Porpidia crustulata</i>	likely	A widespread, Holarctic early colonizer of siliceous pebbles and small stones on the ground. Species not recognizable without microscopic observations
<i>Pyxine soredata</i>	wrong	The image identified as <i>P. soredata</i> refers to another species
<i>P. subcinerea</i>	wrong	The image identified as <i>P. subcinerea</i> refers to another species

Conclusions

iNaturalist is potentially an exceptional tool for citizen science and outreach efforts in lichenology, as shown by the increase in lichen records on the platform. With their year-round occurrence on trees, soil and rocks, and their diverse shapes and colours, lichens are very attractive for projects on biodiversity to counteract 'plant blindness'.

However, due to inherent uncertainties in lichen identification, by untrained or poorly trained volunteers who also mostly cannot access professional tools for correct identification, its application for research purposes is currently limited. While well-defined tasks involving a few identifiable species (e.g. in our study areas *Xanthoria parietina*, *Diploicia canescens*) could be successfully carried out using iNaturalist data, relying on this platform for extensive biodiversity surveys is not recommended. Actually, these activities would increase the risk that some data achieve RG identification and are included in GBIF erroneously. While data collection for large areas with small sampling effort may be useful, it is crucial to generate high-quality data and train users to confidently incorporate observations into research activities.

Acknowledgements. This work received support from the EU through the Erasmus+ Programme and from the grant UIBD/00329/2020 for the Centre for Ecology, Evolution and Environmental Changes (cE3c) from FCT. SR acknowledges the support of NBFC to the University of Palermo, funded by the Italian Ministry of University and Research, PNRR, Missione 4 Componente 2, 'Dalla ricerca all'impresa', Investimento 1.4, Project CN00000033. The authors warmly thank Lucy Sheppard for language revision.

Author ORCIDs.  Silvana Munzi 0000-0002-0071-7699; Deborah Isocrono, 0000-0003-0570-0674; Sonia Ravera 0000-0001-5223-7964.

Competing Interests. The authors declare none.

References

Aristeidou M, Herodotou C, Ballard HL, Young AN, Miller AE, Higgins L and Johnson RF (2021) Exploring the participation of young citizen scientists in scientific research: the case of iNaturalist. *PLoS ONE* **16**, e0245682.

Barve VV, Brenskelle L, Li D, Stucky BJ, Barve NV, Hantak MM, McLean BS, Paluh DJ, Oswald JA, Belitz MW, et al. (2020) Methods for broad-scale plant phenology assessments using citizen scientists' photographs. *Applications in Plant Sciences* **8**, e11315.

Bio Innovation Service (2018) *Citizen science for environmental policy: development of an EU-wide inventory and analysis of selected practices*. Final report for the European Commission, Directorate-General for Environment. In Collaboration with Fundacion I and The Natural History Museum. URL <https://data.europa.eu/doi/10.2779/961304>

Bonney R (1996) Citizen science: a lab tradition. *Living Bird* **15**, 7–15.

Butcher GS, Fuller MR, McAllister LS and Geissler PH (1990) An evaluation of the Christmas Bird Count for monitoring population trends of selected species. *Wildlife Society Bulletin* **18**, 129–134.

Callaghan CT, Poore AG, Hofmann M, Roberts CJ and Pereira HM (2021) Large-bodied birds are over-represented in unstructured citizen science data. *Scientific Reports* **11**, 19073.

Casanovas P, Lynch HJ and Fagan WF (2014) Using citizen science to estimate lichen diversity. *Biological Conservation* **171**, 1–8.

Chozas S, Nunes A, Serrano HC, Ascensão F, Tapia S, Mágues C and Branquinho C (2023) Rescuing Botany: using citizen-science and mobile apps in the classroom and beyond. *npj Biodiversity* **2**, 6.

Clauzade G and Roux C (1985) Likenoj de Okcidenta Europa. Ilustrita determinlibro. *Bulletin de la Société Botanique du Centre-Ouest* **7**, 1–893.

Creley CM and Muchlinski AE (2017) Distribution of the eastern gray squirrel (*Sciurus carolinensis*) within California as of 2015. *Bulletin of the Southern California Academy of Sciences* **116**, 204–213.

de Groot M, Pocock MJ, Bonte J, Fernandez-Conradi P and Valdés-Correcher E (2023) Citizen science and monitoring forest pests: a beneficial alliance? *Current Forestry Reports* **25**, 1–8.

Follett R and Strezov V (2015) An analysis of citizen science-based research: usage and publication patterns. *PLoS ONE* **10**, 1–14.

Hecker S, Garbe L and Bonn A (2018) The European citizen science landscape: a snapshot. In Hecker S, Haklay M, Bowser A, Makuch Z, Voge J and Bonn A (eds), *Citizen Science. Innovation in Open Science, Society and Policy*. London: UCL Press, pp. 190–200.

Hiller T and Haelewaters D (2019) A case of silent invasion: citizen science confirms the presence of *Harmonia axyridis* (Coleoptera, Coccinellidae) in Central America. *PLoS ONE* **14**, e0220082.

Iwanycki Ahlstrand N, Primack RB and Tøttrup AP (2022) A comparison of herbarium and citizen science phenology datasets for detecting response of flowering time to climate change in Denmark. *International Journal of Biometeorology* **66**, 849–862.

- Jackson MM, Gergel SE and Martin K** (2015) Citizen science and field survey observations provide comparable results for mapping Vancouver Island White-tailed Ptarmigan (*Lagopus leucura saxatilis*) distributions. *Biological Conservation* **181**, 162–172.
- Koo KS, Oh JM, Park SJ and Im JY** (2022) Accessing the accuracy of citizen science data based on iNaturalist data. *Diversity* **14**, 316.
- Lee TS, Kahal NL, Kinas HL, Randall LA, Baker TM, Carney VA, Kendell K, Sanderson K and Duke D** (2021) Advancing amphibian conservation through citizen science in urban municipalities. *Diversity* **13**, 211.
- McMullin RT and Allen JL** (2022) An assessment of data accuracy and best practice recommendations for observations of lichens and other taxonomically difficult taxa on iNaturalist. *Botany* **100**, 491–497.
- Nimis PL and Martellos S** (2022) *ITALIC: the information system on Italian lichens. Version 7.0*. Department of Life Sciences, University of Trieste. [WWW document] URL <http://dryades.units.it/italic>. [Accessed 21 January 2023].
- Prudic KL, Oliver JC, Brown BV and Long EC** (2018) Comparisons of citizen science data-gathering approaches to evaluate urban butterfly diversity. *Insects* **9**, 186.
- Ricca PMC and Cheung HCA** (2021) Sighting of the rare jellyfish *Anomalorhiza shawi* Light, 1921 in a marine protected area of Hong Kong. *Check List* **17**, 701–707.
- Rosa RM, Cavallari DC and Salvador RB** (2022) iNaturalist as a tool in the study of tropical molluscs. *PLoS ONE* **17**, e0268048.
- Smith CW, Aptroot A, Coppins BJ, Fletcher A, Gilbert OL, James PW and Wolseley PA** (2009) *The Lichens of Great Britain and Ireland*. London: British Lichen Society.
- Turbé A, Barba J, Pelacho M, Mugdal S, Robinson LD, Serrano-Sanz F, Sanz F, Tsinaraki C, Rubio J-M and Schade S** (2019) Understanding the citizen science landscape for European environmental policy: an assessment and recommendations. *Citizen Science: Theory and Practice* **4**, 34.
- Vardi R, Berger-Tal O and Roll U** (2021) iNaturalist insights illuminate COVID-19 effects on large mammals in urban centers. *Biological Conservation* **254**, 108953.
- Wangyal JT, Bower D, Vernes K and Thinley P** (2022) Employing citizen science to understand amphibian and reptile diversity and distribution in the Himalayan Kingdom of Bhutan. *Global Ecology and Conservation* **37**, e02157.
- Wilson JS, Pan AD, General DE and Koch JB** (2020) More eyes on the prize: an observation of a very rare, threatened species of Philippine Bumble bee, *Bombus irisanensis*, on iNaturalist and the importance of citizen science in conservation biology. *Journal of Insect Conservation* **24**, 727–729.
- Winterton SL** (2020) A new bee-mimicking stiletto fly (*Therevidae*) from China discovered on iNaturalist. *Zootaxa* **4816**, 361–369.