Recognition and completeness: two key metrics for judging the utility of citizen science data

Thomas Mesaglio[1*], Corey T Callaghan[1,2,3,4], Fabrice Samonte[1], Simon BZ Gorta[1], and William K Cornwell[1,5]

Biodiversity citizen science data are being collected at unprecedented scales, and are key for informing conservation and research. Species-level data typically provide the most valuable information, but recognition of specimens to species level from photographs varies among taxa. We examined a large dataset of Australian photographic observations of terrestrial invertebrates uploaded to iNaturalist to quantify recognition to species across different taxa. We also quantified the proportion of Australian species that have been uploaded to iNaturalist. Across 1,013,171 observations covering 14,663 species (17.8% completeness), 617,045 (60.9%) were recognized to species. Dragonflies/damselflies and butterflies were the best-recognized and most complete taxa, and therefore represent the best groups for researchers and managers intending to use existing iNaturalist data at large spatial and temporal scales. The recruitment of additional experts to identify records, and enhanced support for accessible resources for hard-to-identify taxa, will likely increase recognition for other taxa.

Front Ecol Environ 2023; doi:10.1002/fee.2604

Participation in biodiversity-based citizen science (also known as community science) initiatives has increased exponentially over the past several decades (Pocock *et al*. 2017). These initiatives generate data at unprecedented spatial and temporal scales (Dickinson *et al*. 2012), and are invaluable for management, conservation, and monitoring of biodiversity at local to global scales (McKinley *et al*. 2017). Although limitations and biases associated with these data remain (Burgess *et al*. 2017), there are numerous methods that help to account for these constraints, including integrating citizen science data with professional data (Rapacciuolo *et al*. 2021) and filtering or subsampling data (Steen *et al*. 2019). One common bias is that vertebrates and some plant groups are disproportionately common in both professional and citizen science datasets. For example, in biodiversity datasets, birds – many of which are brightly colored and easily detectable and accessible – appear much more frequently (Amano *et al*. 2016) as compared to less charismatic taxa such as invertebrates, mosses, and fungi, which are strongly underrepresented (Troudet *et al*. 2017; Cornwell *et al*. 2019).

Although the contributions of citizen science to invertebrate research have increased markedly in the past few years (Fontaine *et al*. 2021), a major challenge has yet to be overcome: the identification of many organisms to the species level. Many biodiversity citizen science initiatives (eg eBird [https://ebird.org], Reef Life Survey [https://reeflifesurvey.com]) aim to generate data at a species level or lower, and such data typically provide the richest value of information for addressing ecological and conservation research questions. One platform that has already begun to provide reliable species-level data for biodiversity research and conservation is iNaturalist (www.inaturalist.org), an online biodiversity citizen science initiative with almost 2 million users who have contributed >86 million observations (as of December 2021). However, an important consideration is that each datapoint in iNaturalist, and indeed in many modern citizen science projects, is typically a digital photographic record, usually of a living organism, rather than a physical record of a collected and preserved organism, as is traditional for museum and herbarium collections. Here, in a biodiversity data context, we coin the term "recognition" and define it as the proportion of records identified to a species level. We also define recognition for digital records specifically as dependent on a combination of three key factors: (1) the inherent identifiability of an organism, which is influenced by intrinsic qualities (such as organism size, color, and behavior); (2) image quality, which is influenced by extrinsic qualities (such as observer skill and photographic equipment); and (3) the community contribution to identifying records, which is influenced by extrinsic qualities (such as contributor expertise). In addition to variability in recognition between individual organisms, recognition of invertebrates is likely to vary across taxa. Large colorful taxa, such as butterflies or dragonflies, are relatively easy to recognize at the species level, whereas identification of smaller and less conspicuous groups, such as micromoths, often requires genital dissection or even genetic sampling.

[1]*Centre for Ecosystem Science, School of Biological, Earth and Environmental Sciences, The University of New South Wales, Sydney, Australia* [*]*(thomasmesaglio@gmail.com);* [2]*Community Ecology & Conservation Research Group, Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Prague, Czech Republic;* [3]*German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany;* [4]*Department of Wildlife Ecology and Conservation, Fort Lauderdale Research and Education Center, University of Florida, Davie, FL;* [5]*Evolution & Ecology Research Centre, School of Biological, Earth and Environmental Sciences, The University of New South Wales, Sydney, Australia*

Along with recognition, and also important for biodiversity data and research, is "completeness", which we define here as the proportion of all known, described species in a defined area that have been observed in a given dataset. Recognition and completeness are linked to some degree, in that a species cannot contribute to the completeness of a dataset without first being recognized. As such, the probability of any given species being observed is unequal, and is influenced by the same key factors outlined above. To our knowledge, recognition and completeness across invertebrate taxa from citizen science photographs have not been previously quantified.

We examined a dataset of >1,000,000 Australian photographic observations of terrestrial and semi-aquatic invertebrates uploaded to iNaturalist to document when and where limits to species recognition exist. We also quantified the completeness of different taxa (ie the proportion of all described Australian species per taxon that have been uploaded to iNaturalist at least once). To understand trends in time and space in recognition and completeness, we compared these data at two time points, and with data from two additional regions: the Netherlands and Taiwan.

Methods

On two separate occasions, we extracted all georeferenced and dated records associated with a photograph within continental Australia and islands under Australian jurisdiction from iNaturalist for four invertebrate phyla with terrestrial representatives and at least 100 Australian observations (Annelida, Arthropoda, Mollusca, Platyhelminthes). The first dataset contained observations that had been uploaded on or before 5 November 2020 (n = 527,313), and the second dataset contained observations uploaded on or before 5 December 2021 (n = 1,013,171). We categorized the records within each dataset into 39 "iconic" taxa (WebTable 1): that is, recognizable groups (eg "butterflies", "ants") ranging from subclass to family (see WebPanel 1 for full details about data filtering and processing). For spatial comparisons with Australia, we repeated this process for two distinctly different regions of the world: the Netherlands and Taiwan (WebPanel 1). These regions were chosen because they represent different biodiversity levels, climate zones, iNaturalist use, and languages, and because a comprehensive species checklist was available for both.

For each iconic taxon, we calculated the proportion of records identified to species level (recognition). We also calculated the proportion of described species within the selected taxa that have been identified on the iNaturalist platform (completeness) by comparing species counts with known published biodiversity datasets (see WebPanel 1). For the 2021 Australia dataset, recognition was also calculated for the five most observed families within four of the most observed insect orders to highlight how recognition differs within iconic taxa.

For the 2021 Australia, Netherlands, and Taiwan datasets, we calculated species evenness (J') for each taxon using Pielou's evenness index (Pielou 1966) to quantify variability in recognition between species (see WebPanel 1). For iconic taxa with low evenness values, the distribution of observations identified to species is more biased toward a few common species. We also determined correlations between the 2021 Australia, Netherlands, and Taiwan datasets for recognition, completeness, and species evenness (R Core Team 2021).

Results

As of 5 December 2021, 617,045 of 1,013,171 (60.9%) Australian observations of terrestrial invertebrates were recognized to species. When grouped into 39 iconic taxa, the average recognition of a taxon was 44.4%, ranging from 0.4% (mayflies) to 97.2% (dragonflies/damselflies) (Figure 1). Of the 81,956 described Australian species within the taxa we explored, 14,663 (17.9%) had been observed on iNaturalist. Across our 39 taxa, average completeness was 24.2%, ranging from 0.4% (pseudoscorpions and short-tailed whipscorpions) to 81.9% (dragonflies/damselflies). The average species evenness across all iconic taxa was 0.656 (WebTable 2), ranging from 0.249 (millipedes, low evenness) to 1 (mayflies, high evenness). Recognition also differed across taxonomic levels within many iconic taxa, with poorly recognized taxa still containing families or other finer-level taxa with high recognition, and vice versa (Figure 2).

For 38 of the 39 iconic taxa within the two Australian datasets, from 5 November 2020 to 5 December 2021, completeness increased, ranging from +0.23% (bark, book, and true lice) to +12.12% (scorpionflies); for the remaining taxon (pseudoscorpions), completeness did not change. During this time period, recognition increased for 26 taxa and decreased for 13 taxa, ranging from −6.17% (thrips) to +15.73% (dobsonflies, alderflies, and allies). For all taxa, the number of observations increased, ranging from a 1.70-fold increase (moths) to a 2.38-fold increase (earthworms and leeches). Indeed, for 23 taxa, the number of observations more than doubled. Slugs and land snails (+92 species), hoppers, scale insects, and allies (+142), and mites and ticks (+12) all increased in number of species by almost 50%, whereas flies (+277), moths (+636), and beetles (+794) all increased by more than 250 species.

For the Netherlands dataset, the average recognition of a taxon was 65.7% and the average completeness of a taxon was 32.8%; for the Taiwan dataset, the average recognition of a taxon was 39.2% and the average completeness of a taxon was 35.2% (WebPanel 2; WebFigure 1).

Across the iconic taxa, completeness was strongly correlated between the 2021 Australia dataset, Netherlands dataset, and Taiwan dataset. For recognition and species evenness across the iconic taxa, the Australia and Taiwan datasets were strongly correlated, but the correlation was weak between the Netherlands dataset and each of the Australia and Taiwan datasets (WebTable 3).

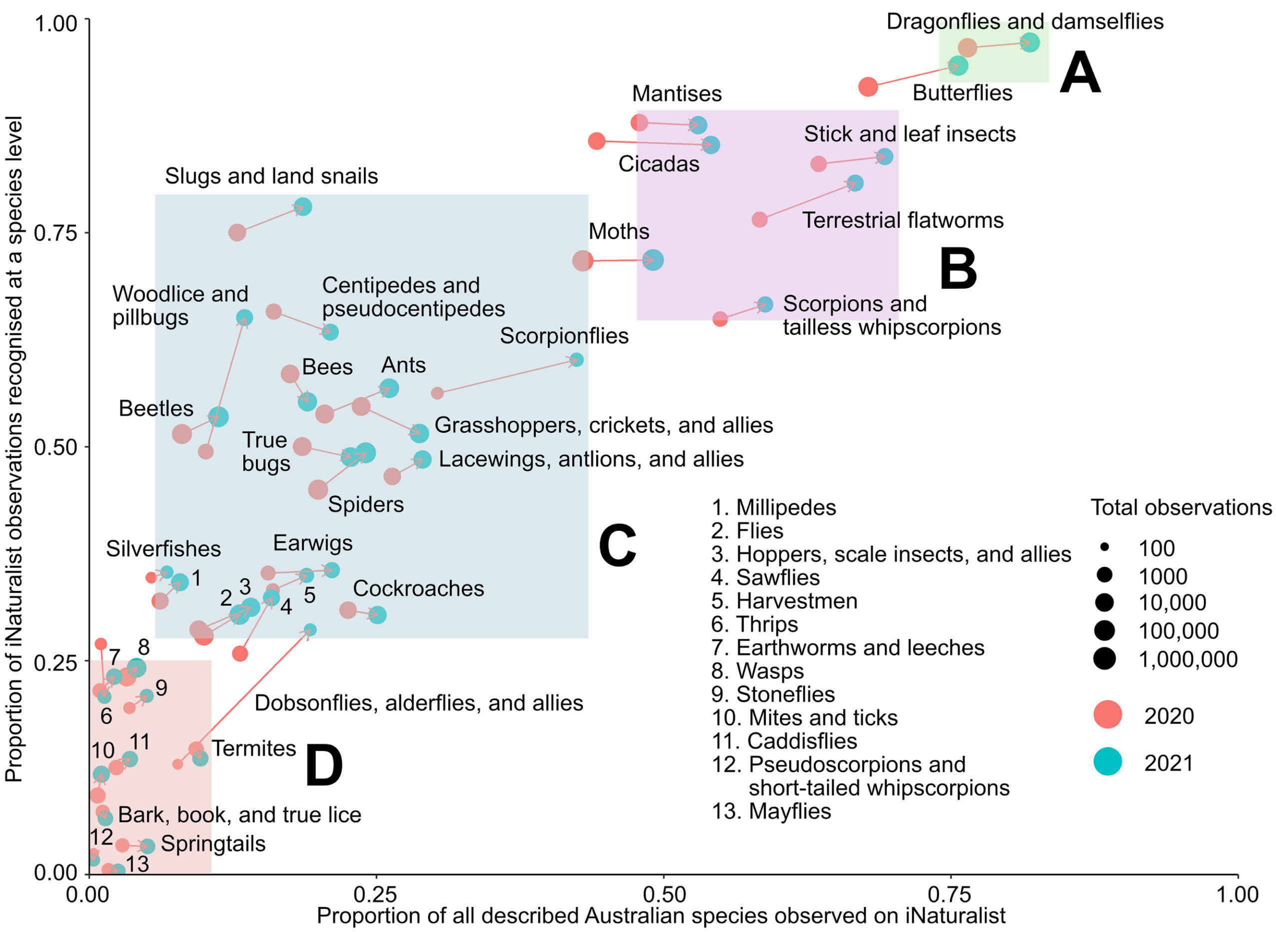

Figure 1. Recognition and completeness of Australian terrestrial invertebrates from photographs uploaded to iNaturalist across 39 iconic taxa (for 2020, n = 527,313; for 2021, n = 1,013,171). The size of each point is scaled to the log of the number of observations for that taxon. The colored rectangles were manually drawn to group what we perceived as taxa with similar recognition and completeness on iNaturalist, based on the 2021 dataset. Capital letters A–D denote each group.

Discussion

For the 2021 Australia dataset, we categorized the iconic taxa into four discrete groups (Figure 1) based on recognition and completeness using the following thresholds: Group A, >90% recognition and >75% completeness; Group B, 70–90% recognition and >45% completeness; Group C, >25% recognition and 5–45% completeness; and Group D, 0–25% recognition and 0–10% completeness. It is important to note that there was no statistical basis for these groups, and that recognition and completeness fall along a continuous gradient; however, categorization allows for a clearer discussion of the differences among taxa, and thus we created the four groups to aid our discussion and recommendations.

Recognition

For Australian terrestrial invertebrates, recognition based on photographs uploaded to iNaturalist varied greatly among higher-level taxa, ranging from under 1% to almost 100%. Only two taxa above the family level had over 90% of observations recognized at the species level: butterflies (94.5%) and dragonflies/damselflies (97.2%). These two taxa share six key aspects (categorized under the three factors we referred to earlier) relating to recognition: most or many species (1) are relatively large and easy to photograph; (2) have a consistent body plan and general shape; (3) are colorful or well-patterned; and (4) are recognized based on colors or patterning (ie characters typically readily visible even in low-quality photographs); in addition, both groups are associated with (5) a number of Australian experts who are regular users of iNaturalist and (6) comprehensive, popular field guides accessible to non-experts. This level of recognition allows for powerful insights into species trends at large spatial and temporal scales (Forister *et al.* 2021). Expanding this recognition level to other taxa is a crucial goal. Although we included both "needs ID" and "Research Grade" observations (see Mesaglio and Callaghan 2021) in our datasets, the former are not necessarily less taxonomically accurate than the latter (Hochmair *et al.* 2020). We note that in some cases observations may be misidentified as an incorrect species, which can impact recognition; however, such misidentifications are not typically of "new"

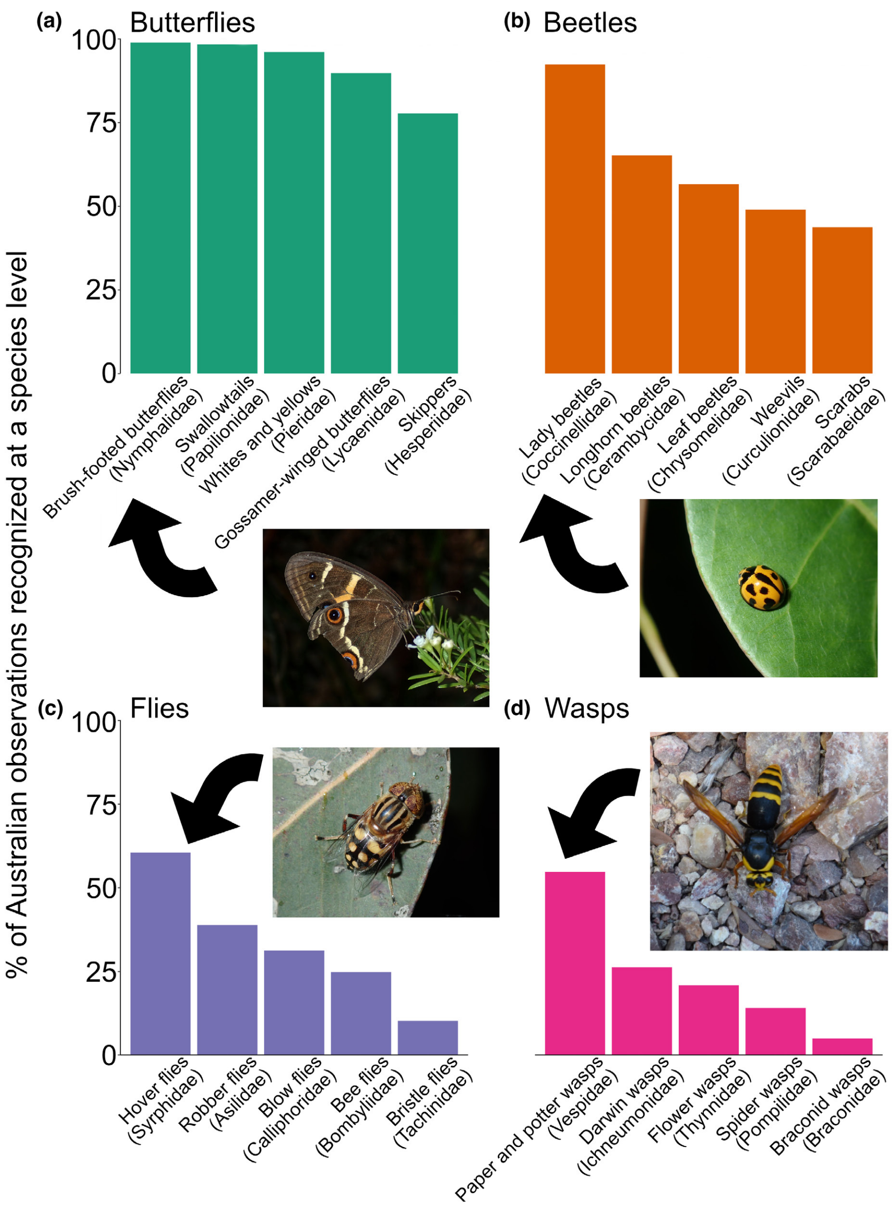

Figure 2. Recognition across the top five most observed families within four of the most observed insect groups on iNaturalist Australia. (a) swordgrass brown (*Tisiphone abeona*), (b) variable ladybird (*Coelophora inaequalis*), (c) native drone fly (*Eristalinus punctulatus*), and (d) *Rhynchium superbum*. All photos by T Mesaglio.

species for iNaturalist, and therefore have little effect on completion.

These six factors have unequal influences on recognition; expert engagement with citizen science can have a disproportionate impact when the important traits for recognition are easily photographed without specific training. For example, despite the small size and typically dull coloration of many species, land snails and slugs are well recognized (78.1% of observations identified to species). Although this high percentage is in part driven by many observations of easily recognized nonnative species (eg garden snails [*Cornu aspersum*] and leopard slugs [*Limax maximus*]), it is also driven by the high activity of one particular Australian expert, Kevin Bonham, who has provided identifications on almost 90% of all Australian observations of land snails and slugs. A similar case exists for terrestrial flatworms (for which 80.8% of observations have been identified to species), with Australian and international authority Leigh Winsor also a regular identifier on iNaturalist, and many other finer-level taxa as well (see WebPanel 3), highlighting the disproportionate influence of even a single expert identifier. However, some of the taxa in Group C and many in Group D are unlikely to greatly benefit from increased expert participation alone. For instance, millipede identification is largely based on microscopic characteristics, which are not captured in most citizen science photographs. Interestingly, the position of millipedes in Group C is strongly driven by the large number of observations of the nonnative Portuguese millipede (*Ommatoiulus moreleti*); removing this species shifts millipedes from 34.2% recognition to 9.4% and into Group D, underscoring the importance of accounting for the disproportionate influence of a few easily recognized and well-observed species in some taxa. This pattern was also present in other taxa in which there are one or two very common and easily recognized nonnative species, such as bees (*Apis mellifera*), silverfishes (*Ctenolepisma longicaudata*), and woodlice and pillbugs (*Armadillidium vulgare*, *Porcellio scaber*), compared to many native species that are more difficult to recognize to species. Cases like these also indicate that, although recognition and completeness are inherently linked, this relatedness is not an issue for our framework. This is highlighted by the variable evenness indices across our iconic taxa, with an unequal probability of each species being observed (WebTable 2).

Completeness

As with recognition, the proportion of all described Australian species uploaded to iNaturalist is highly variable across higher-level taxa, ranging from under 1% to over 80%. Not surprisingly, the most complete groups for taxa above the family level – dragonflies/damselflies (81.9%) and butterflies (75.6%) – have many large, active, and colorful species, an observation bias reported for both citizen science and professional insect monitoring, and both within and between higher taxa (Dennis *et al.* 2006; Ward 2014; Lobo *et al.* 2021). These taxa also have relatively low diversity (331 and 447 species, respectively) compared to taxa such as beetles (2813 species uploaded to iNaturalist out of 25,017 described Australian species) or wasps (361 uploaded out of 8840 described).

Moving from Group A downward, and particularly within Groups C and D, taxa tend to increasingly become less charismatic, be more diverse, exhibit more cryptic behavior, have more specific habitat requirements, and have smaller spatial ranges, all of which contribute to reduced likelihood of detection. Given that the addition of new species to iNaturalist in Australia is increasing at an exponential rate (Mesaglio and Callaghan 2021), completeness for many currently underrepresented invertebrate taxa will naturally increase over time – which, as noted above, was the case for 38 of the 39 iconic taxa in the 13 months between 5 November 2020 and 5 December 2021. However, there are also important opportunities to make biodiversity data more complete by actively encouraging increased observation of small, cryptic taxa. Moreover, many "unobserved" species have indeed already been observed and uploaded to iNaturalist but have not yet been recognized. Greater recruitment of experts will therefore improve both recognition and completeness for many taxa. A key future research avenue to expand our framework will be to understand how species traits may help explain variability in recognition.

Comparison between 2020 and 2021 data

Encouragingly, in the 13 months between 5 November 2020 and 5 December 2021, the number of observations increased between 1.7-fold and 2.38-fold among our iconic taxa, consistent with growing participation in citizen science (Mesaglio and Callaghan 2021). Completeness, number of observations, and number of species increased for all taxa except pseudoscorpions. Recognition increased for 26 of 39 taxa, and of the 13 taxa for which recognition decreased, five were in Group C and six were in Group D. We suspect that increased recognition reflects greater expert engagement with the community (see WebPanel 3) while declines occur when experts are absent.

Comparison between regions

We found the often-high variability in recognition and completeness across taxa between different regions (Australia, the Netherlands, and Taiwan) to be largely due to differences in local diversity. For example, for Australia and Taiwan, there are observations of 14 and ten scorpionfly species respectively, exceeding the six species observed for the Netherlands; however, this represents 100% completion for the Netherlands, whereas completion is considerably lower for Australia (42%) and Taiwan (21.3%). Although this is a more extreme example, with completion relatively strongly correlated between the three datasets (WebTable 3), recognition does vary considerably; average recognition for the Netherlands (65.7%) far exceeded that for Australia (44.4%) and Taiwan (39.2%), a product of the Netherlands' lower diversity across the focal taxa and because this diversity is better understood (eg fewer undescribed species, more identification resources). Therefore, for relatively low diversity and well-studied regions, the thresholds for completion and recognition for Groups A and B are likely easier to reach, and greater attention can be directed toward increasing the total number of observations (the Netherlands dataset comprised just 10.5% of the number of observations in the Australia dataset).

Use of data for research and conservation

Currently, most structured research using iNaturalist data is performed with datasets that involve taxa from our Group A, consisting of butterflies (eg Forister *et al.* 2021) and dragonflies/damselflies (eg Drury *et al.* 2019; Bowler *et al.* 2021). Accounting for both recognition and completeness, these two taxa were also the best performers in the Netherlands and Taiwan datasets. The dominance of butterflies and dragonflies/damselflies in citizen science is also echoed in professional science and conservation (Cowie *et al.* 2022). Butterflies are disproportionately common as indicator taxa in studies of environmental change (Fleishman and Murphy 2009), and of the 123 protected European insect species as of 2017, ~42% were butterflies, dragonflies, and damselflies (Leandro *et al.* 2017).

On the basis of recognition and completeness, we therefore recommend that, for researchers and managers who

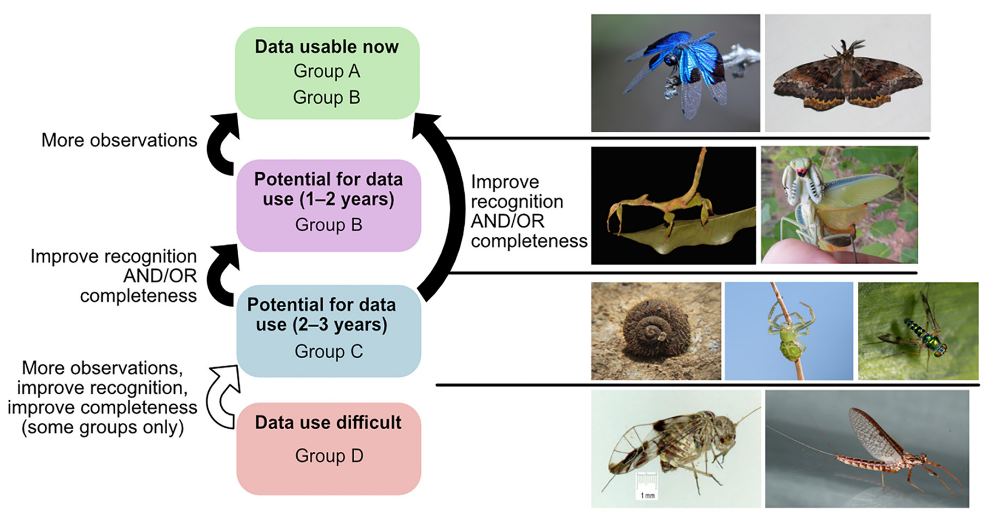

Figure 3. Categorization of Groups A–D from Figure 1 into four levels based on how "usable" the data are for research and conservation. Starting at the top and moving from left to right, first row: jewel flutterer (*Rhyothemis resplendens*), G Winterflood (@graham_winterflood); batwing moth (*Chelepteryx collesi*), J Lenagan (@johnlenagan). Second row: spiny leaf insect (*Extatosoma tiaratum*), N Lambert (@nicklambert); rainbow mantis (*Sphodropoda quinquedens*), M Connors (@matthew_connors). Third row: *Tolgachloritis campbelli*, M Connors (@matthew_connors); undescribed crab spider (Thomisidae), R Richter (@reiner); *Austrosciapus* sp., M Ayers (@markayers). Fourth row: unidentified barklouse (Philotarsidae), G Cocks (@graemevc); unidentified mayfly (Ephemeroptera), V Fazio III (@vicfazio3). iNaturalist usernames appear in parentheses after photographers' names.

intend to use existing iNaturalist data (whether from Australia or elsewhere) at large spatial and temporal scales, butterflies and dragonflies/damselflies are currently the best options (Figure 3), especially for diversity-related analyses or in visualization of spatiotemporal trends (similar to, for example, Forister *et al.* 2021). For instance, butterflies and dragonflies/damselflies would be strong candidates for species-level distribution modeling given the intense sampling effort applied to these taxa, and indeed Australian iNaturalist data have already been used to help map the shifting distribution of the tawny coster (*Acraea terpsicore*) (Chowdhury *et al.* 2021), a butterfly species that has spread from South Asia into northern Australia. These taxa are also strong candidates for estimating distribution trends using occupancy models (Bowler *et al.* 2021), and opportunistic butterfly and dragonfly/damselfly data have already been used to estimate large-scale distribution trends when analyzed using occupancy models and integrated with structured monitoring data (van Strien *et al.* 2013). Analyses such as these highlight the importance of recognition – if recognition is low for a given species, there will not be enough data for distribution models. The greater the minimum sample size (that is, the higher the recognition), the more robust these models will be.

Despite having lower recognition and completeness as compared to Group A, moths are also a strong candidate for research using existing iNaturalist data. Moths comprise 12% of all Australian iNaturalist observations across all taxa and 34.4% of all observations in our Australia dataset. The relatively low completeness of described Australian moth species uploaded to iNaturalist is a product of their high diversity (>10,000 species), and their absolute diversity on iNaturalist far exceeds that of any other terrestrial invertebrate taxon. Moths also have the largest number of species, and the highest or second highest number of observations, in both the Netherlands and Taiwan datasets. We therefore highlight moths as a third taxon of high potential for insightful analyses in the short-to-medium term in Australia and other regions of the world.

For groups with lower recognition and completeness, there are opportunities for improvement through increased data collection. Group B taxa are the most promising based on their already relatively high recognition (>70%) and completeness (>45%), with potential to reach the current level of Group A and be useful for future research applications within the next few years. Indeed, we believe most of the taxa in Group B already have usable data from a recognition and completeness perspective (given that mantises, cicadas, stick/leaf insects, and terrestrial flatworms all have recognition exceeding 80%), and that the most important factor will be promoting more observations of these taxa given that they are relatively undersampled. This offers the potential for important discoveries because much less is known about the ecology of Group B taxa as compared to Group A taxa, including from professionally collected datasets.

Many of the Group C taxa also show potential to move into Group B, albeit with more time needed not only for recognition and completeness to improve (both organically and with greater recruitment of taxonomic experts) but also for the number of observations to increase. Several Group C taxa have the potential to move directly into Group A through improvements to recognition and completeness given their already large number of observations; for example, spiders, beetles, and flies all have more observations than butterflies or dragonflies/damselflies. Many families or genera within our iconic taxa already have sufficient data to allow for distribution modeling and other applications. For instance, although ants (family Formicidae) are placed in our Group C, the genus *Myrmecia* (bull ants) has 83.5% recognition, 60.9% completion, and >6000 observations, placing it toward the upper end of Group B. However, for many Group D taxa, such as mites and ticks or thrips, the resolution of citizen science photographs is often too low to allow identification, regardless of the geographic region of focus and even if assessed by an expert; therefore, improving image quality is one of the key factors for improving recognition and completeness. It is important to acknowledge that there are still opportunities for data use in any of our iconic taxa, regardless of their completeness or recognition, especially for analyses focusing on just one or two species. For example, despite the very low completeness of millipedes (7.9%), there are 1623 observations of the Portuguese millipede; these data could be used to track spatiotemporal trends in the spread of this species across Australia. Also importantly, we highlight that our framework is most useful for analyses of broader-level taxa such as orders or families. For studies focusing on only a few indicator species within specific groups, number of observations is a more relevant metric than recognition or completeness. Such studies allow taxa with relatively low overall completeness to still provide usable data, provided those few species have large sample sizes.

Our framework may also provide a valuable recommendatory tool for conservation. Under many national and international conservation frameworks, taxa coarser than species are not assigned conservation statuses. This is problematic for taxa where recognition to species level is difficult (whether in professional or citizen science), leading to insufficient distribution, abundance, and trend data, which may prevent potential listing (Marsh *et al.* 2021). The recognition of higher-level taxa (eg genera) as having persistently low recognition despite expert engagement may provide an important line of evidence for affording these taxa formal protection. For example, Australian mayflies are very difficult to recognize to species level from photographs (0.4% recognition for our 2021 Australian dataset). New, identified, collected material is rare; zero Australian mayfly records were added to the Global Biodiversity Information Facility between 2020 and 2021, whereas 552 observations were added to iNaturalist over this same time period, suggesting that detecting any near-term declines in mayfly taxa would require the use of photographic data (see Forister *et al.* 2021). However, because species-level recognition is low, this precludes species-level action. Focusing on higher-level taxa, such as genera or families, offers a way forward.

Conclusions

In light of the growing importance of citizen science data (Chandler *et al*. 2017; Forister *et al*. 2021), it is imperative that efforts be made to improve these data through increasing recognition, completeness, and number of observations. We suggest that one way in which professional scientists and funding bodies can achieve this is through funding and supporting better and more accessible identification resources, such as those that currently exist for butterflies and dragonflies/damselflies in some regions. Also key is boosting incentives for experts to contribute their time to identifying observations. Increasingly, institutions are recognizing identifications made on iNaturalist as scientific outreach and including them as part of paid curatorial time and within performance reviews. Greater adoption of these measures by greater numbers of scientific organizations will likely be an important driving force for increasing expert contributions. Given the exponentially expanding efforts of citizen scientists (Mesaglio and Callaghan 2021; Ruiz-Gutierrez *et al*. 2021), investing in better resources will generate high return-on-investment in terms of monitoring data, including for such sectors as biosecurity (Thomas *et al*. 2017). Moreover, this process will expand our knowledge base about population trends beyond a few charismatic taxa and allow for a more taxonomically robust understanding of biodiversity.

Acknowledgements

We thank Bob Mesibov for invaluable insights and suggestions, especially regarding the concept of "recognition", and Hayden Schilling for helpful comments.

Data Availability Statement

Our code is not novel. Data and code are available at https://github.com/tmesaglio/iNaturalist-invertebrates-recognition.

References

Amano T, Lamming JD, and Sutherland WJ. 2016. Spatial gaps in global biodiversity information and the role of citizen science. *BioScience* **66**: 393–400.

Bowler DE, Eichenberg D, Conze KJ, *et al*. 2021. Winners and losers over 35 years of dragonfly and damselfly distributional change in Germany. *Divers Distrib* **27**: 1353–66.

Burgess HK, DeBey LB, Froehlich HE, *et al*. 2017. The science of citizen science: exploring barriers to use as a primary research tool. *Biol Conserv* **208**: 113–20.

Chandler M, See L, Copas K, *et al*. 2017. Contribution of citizen science towards international biodiversity monitoring. *Biol Conserv* **213**: 280–94.

Chowdhury S, Braby MF, Fuller RA, and Zalucki MP. 2021. Coasting along to a wider range: niche conservatism in the recent range expansion of the tawny coster, *Acraea terpsicore* (Lepidoptera: Nymphalidae). *Divers Distrib* **27**: 402–15.

Cornwell WK, Pearse WD, Dalrymple RL, and Zanne AE. 2019. What we (don't) know about global plant diversity. *Ecography* **42**: 1819–31.

Cowie RH, Bouchet P, and Fontaine B. 2022. The sixth mass extinction: fact, fiction or speculation? *Biol Rev* **97**: 640–63.

Dennis RLH, Shreeve TG, Isaac NJB, *et al*. 2006. The effects of visual apparency on bias in butterfly recording and monitoring. *Biol Conserv* **128**: 486–92.

Dickinson JL, Shirk J, Bonter D, *et al*. 2012. The current state of citizen science as a tool for ecological research and public engagement. *Front Ecol Environ* **10**: 291–97.

Drury JP, Barnes M, Finneran AE, *et al*. 2019. Continent-scale phenotype mapping using citizen scientists' photographs. *Ecography* **42**: 1436–45.

Fleishman E and Murphy DD. 2009. A realistic assessment of the indicator potential of butterflies and other charismatic taxonomic groups. *Conserv Biol* **23**: 1109–16.

Fontaine C, Fontaine B, and Prévot AC. 2021. Do amateurs and citizen science fill the gaps left by scientists? *Curr Opin Insect Sci* **46**: 83–87.

Forister ML, Halsch CA, Nice CC, *et al*. 2021. Fewer butterflies seen by community scientists across the warming and drying landscapes of the American West. *Science* **371**: 1042–45.

Hochmair HH, Scheffrahn RH, Basille M, and Boone M. 2020. Evaluating the data quality of iNaturalist termite records. *PLoS ONE* **15**: e0226534.

Leandro C, Jay-Robert P, and Vergnes A. 2017. Bias and perspectives in insect conservation: a European scale analysis. *Biol Conserv* **215**: 213–24.

Lobo JM, Gutiérrez-Cánovas C, Fox R, *et al*. 2021. Species traits influence the process of biodiversity inventorying: a case study using the British butterfly database. *Insect Conserv Diver* **14**: 748–55.

Marsh J, Bal P, Fraser H, *et al*. 2021. Assessment of the impacts of the 2019–20 wildfires of southern and eastern Australia on invertebrate species. Brisbane, Australia: National Environmental Science Programme.

McKinley DC, Miller-Rushing AJ, Ballard HL, *et al*. 2017. Citizen science can improve conservation science, natural resource management, and environmental protection. *Biol Conserv* **208**: 15–28.

Mesaglio T and Callaghan CT. 2021. An overview of the history, current contributions and future outlook of iNaturalist in Australia. *Wildlife Res* **48**: 289–303.

Pielou EC. 1966. The measurement of diversity in different types of biological collections. *J Theor Biol* **13**: 131–44.

Pocock MJ, Tweddle JC, Savage J, *et al*. 2017. The diversity and evolution of ecological and environmental citizen science. *PLoS ONE* **12**: e0172579.

R Core Team. 2021. R: a language and environment for statistical computing. Vienna, Austria: The R Foundation for Statistical Computing.

Rapacciuolo G, Young A, and Johnson R. 2021. Deriving indicators of biodiversity change from unstructured community-contributed data. *Oikos* **130**: 1225–39.

Ruiz-Gutierrez V, Bjerre ER, Otto MC, *et al*. 2021. A pathway for citizen science data to inform policy: a case study using eBird data for defining low-risk collision areas for wind energy development. *J Appl Ecol* **58**: 1104–11.

Steen VA, Elphick CS, and Tingley MW. 2019. An evaluation of stringent filtering to improve species distribution models from citizen science data. *Divers Distrib* **25**: 1857–69.

Thomas ML, Gunawardene N, Horton K, *et al*. 2017. Many eyes on the ground: citizen science is an effective early detection tool for biosecurity. *Biol Invasions* **19**: 2751–65.

Troudet J, Grandcolas P, Blin A, *et al*. 2017. Taxonomic bias in biodiversity data and societal preferences. *Sci Rep-UK* **7**: 9132.

van Strien AJ, van Swaay CA, and Termaat T. 2013. Opportunistic citizen science data of animal species produce reliable estimates of distribution trends if analysed with occupancy models. *J Appl Ecol* **50**: 1450–58.

Ward DF. 2014. Understanding sampling and taxonomic biases recorded by citizen scientists. *J Insect Conserv* **18**: 753–56.

Supporting Information

Additional, web-only material may be found in the online version of this article at http://onlinelibrary.wiley.com/doi/10.1002/fee.2604/suppinfo