nature ecology & evolution

PERSPECTIVE

https://doi.org/10.1038/s41559-018-0608-1

A decision tree for assessing the risks and benefits of publishing biodiversity data

Ayesha I. T. Tulloch[1,2,3*], Nancy Auerbach[4], Stephanie Avery-Gomm[4,5], Elisa Bayraktarov[4], Nathalie Butt[4], Chris R. Dickman[3], Glenn Ehmke[6], Diana O. Fisher[4], Hedley Grantham[2], Matthew H. Holden[5], Tyrone H. Lavery[7], Nicholas P. Leseberg[1], Miles Nicholls[8], James O'Connor[6], Leslie Roberson[4], Anita K. Smyth[9], Zoe Stone[1], Vivitskaia Tulloch[4], Eren Turak[10,11], Glenda M. Wardle[3] and James E. M. Watson[1,2]

Inadequate information on the geographical distribution of biodiversity hampers decision-making for conservation. Major efforts are underway to fill knowledge gaps, but there are increasing concerns that publishing the locations of species is dangerous, particularly for species at risk of exploitation. While we recognize that well-informed control of location data for highly sensitive taxa is necessary to avoid risks, such as poaching or habitat disturbance by recreational visitors, we argue that ignoring the benefits of sharing biodiversity data could unnecessarily obstruct conservation efforts for species and locations with low risks of exploitation. We provide a decision tree protocol for scientists that systematically considers both the risks of exploitation and potential benefits of increased conservation activities. Our protocol helps scientists assess the impacts of publishing biodiversity data and aims to enhance conservation opportunities, promote community engagement and reduce duplication of survey efforts.

Achieving effective conservation relies on accurate knowledge of where species occur to assist with their management[1–3]. This is particularly true for rare and endangered species that are at risk of extinction. Despite this, one in six International Union for Conservation of Nature (IUCN)-listed species are considered data deficient, and conservation practitioners routinely face a paucity of primary data on the temporal and spatial distribution of biodiversity[4–6]. Resolving this issue is urgent: without adequate spatially explicit biodiversity data, good management and policy decisions that enable the protection of species and ecosystems may be unachievable[7–9].

Primary biodiversity data are evidence that associates a species or taxon with a geographic location within a specified time interval. This may include one or more types of evidence: a sighting; a DNA sample; a verified photographic image; or traces such as scats, tracks, nests or burrows that that can be attributed to a given taxon with confidence. Primary data may also provide biologically useful information such as age, sex, breeding status and population abundance. Today there are not just unprecedented online science data services for researchers, conservationists and the public (for example, wildlife atlases and scientific data repositories such as http://aekos.org.au)[10], but an increased willingness to share primary biodiversity data (for example, via citizen science programmes such as eBird)[11]. Furthermore, scientific journals and funding agencies increasingly request transparent archiving of research data[12–14].

Sharing species occurrence information publicly or privately presents a challenge for scientists because it requires balancing potentially difficult and uncertain trade-offs. For example, shortly after their discovery was published, poaching for the pet trade contributed to the local extinction of Chinese cave geckos *Goniurosaurus luii* in Vietnam[15], prompting calls not to publish primary biodiversity data[16]. In contrast, primary occurrence data shared by researchers in publicly available databases and within the scientific literature were critical to recent re-assessments of extinction risk for endemic birds in Bolivia and Australia[17,18], which allowed for accurate assessments of extinction status of up to two-thirds of the examined species that otherwise would have been uncertain. To ensure effective conservation informed by the best available knowledge of species distributions and abundances, we must understand the benefits of sharing data and the costs of not sharing data, rather than only the risks as has been the recent focus. Here, we propose a risk management decision protocol that balances potential negative outcomes for species against the conservation benefits of publishing primary occurrence data. By following our decision tree, scientists collecting biodiversity data will be able to ensure that they do not overlook potential conservation opportunities for study species, and that conservation mistakes do not occur through inappropriate release or restriction of data.

How biodiversity data are shared

Data publication is often carefully managed by data authors and custodians to maintain confidentiality and meet jurisdictional laws

[1]School of Earth and Environmental Sciences, The University of Queensland, St. Lucia, Queensland, Australia. [2]Wildlife Conservation Society, Global Conservation Program, Bronx, NY, USA. [3]Desert Ecology Research Group, School of Life and Environmental Sciences, University of Sydney, Sydney, New South Wales, Australia. [4]School of Biological Sciences, The University of Queensland, St. Lucia, Queensland, Australia. [5]Australian Research Council Centre of Excellence for Environmental Decisions, University of Queensland, St. Lucia, Queensland, Australia. [6]BirdLife Australia, Melbourne, Victoria, Australia. [7]Biodiversity Institute, University of Kansas, Lawrence, KS, USA. [8]Atlas of Living Australia, National Collections & Marine Infrastructure, CSIRO, Acton, Australian Capital Territory, Australia. [9]Terrestrial Ecosystem Research Network (TERN), The University of Adelaide, Adelaide, South Australia, Australia. [10]NSW Office of Environment and Heritage, Sydney, New South Wales, Australia. [11]Australian Museum, Sydney, New South Wales, Australia. *e-mail: ayesha.tulloch@sydney.edu.au

Table 1 | Likely objectives of various data owners (potential data providers) for sharing biodiversity data and perceptions of how publishing species data might advance or undermine that objective

Objective for sharing data	Examples of data authors	Perception of how publishing data advances objective	Perception of how publishing data undermines objective
Minimize risk of species extinction	Government agencies, NGO ecologists, scientists, land managers	Avoid impacts and increase protection in known locations; increase species ecological knowledge; inform status assessments (for example, IUCN Red List)	Increase human access to species locations, leading to poaching[51], disease spread, habitat degradation (trampling, alien species spread) or disturbance[36,52]
Maximize ability to get funding for biodiversity management	NGO ecologists, scientists, land managers	Secure future funding by demonstrating importance and effectiveness of project (for example, important species detected)	Reduce future funding by disproving importance or effectiveness of project (for example, important species not detected), or exposing inappropriately collected data; hinder projects and profitability by inspiring extra regulation
Inform and improve future monitoring	Government agencies at local, state or national levels	Evaluate progress towards conservation objectives (for example, identify trends); develop spatial units for data aggregation and future sampling	Violate privacy of stakeholders granting access to study area (for example, landholders, indigenous groups), limiting future access
Maximize data reliability for improving knowledge	IUCN, government agencies, scientists	Decrease false negative error rate; illuminate gaps in knowledge of species; reduce data deficiency for species[53]	Increase false positive error rate if data sets vary in rigour or accuracy, wasting effort on incorrect records
Minimize data loss/entropy and survey redundancy	Museums, developers, consultants, government agencies	Reduce risk of losing access to data and legacy effects of loss of local champion[54]; maximize efficiency of limited resources to collect structured data	Increase visitation to private study or recreational locations, reducing novelty or utility of data; reduce knowledge power and competitive edge (for example, to win survey tenders through local expertise)
Maximize academic output or impact	Scientific researcher in academic institution	Increase collaborations, citations and use of research[55]; verify published results	Others use data without acknowledging/citing source[56,57]; opportunity cost (for example, time) of compiling/sharing data; embarrassment if others find errors in analysis (for example, publish rebuttals)
Maximize public interest in biodiversity	Citizen science programmes, citizen scientists, government agencies	Engage the public in nature; human health benefits of wildlife access[58]	Increase human access to species locations, leading to inadvertent habitat degradation or disturbance and reduced site value[36,52]

and national regulations (Supplementary Table 1). Ways of managing the release of data classified as 'sensitive' range from publishing precise locations but changing species identifiers to a classification of 'restricted' or to a higher taxonomic resolution such as genus or family (if spatial locations are important to share for conservation purposes), to keeping species names accurate but changing locations to mask true spatial coordinates (for example, by buffering or masking the location), or restricting species location information completely by withholding it from public access (see Supplementary Table 1).

The most comprehensive guide on assessing sensitivities around species and required generalization rules for publishing species locations is provided by the Global Biodiversity Information Facility (GBIF)[19]. GBIF's protocol is first to identify which species are at risk from harm by human activity, and second to assess the impact of this activity on the taxon. These criteria are used to determine whether a species is flagged as sensitive and are then followed by further rules determining the degree of sensitivity. A subsequent rule determines whether release of information will increase the likelihood of harmful impacts on the species. The assessment for whether data should be released considers what level of generalization or 'denaturing' might be required. These range from no restriction for species classed as 'low sensitivity', to increasing restrictions through data generalization for 'low to medium sensitivity' (0.001°), 'medium to high sensitivity' (0.01°) and 'highly sensitive taxa' (0.1°). All location data are withheld if a species is identified as being of high biological significance and under high threat[19]. However, no consideration of the benefits of publishing data is made.

There are methods of publishing information on where species occur that do not directly release raw species locality data. Many non-governmental organization (NGO) expeditions assess and publish data on the biological value of areas to highlight the need for conservation action; for example, Conservation International's Rapid Assessment Program shares expedition data online to promote awareness of regions with high biodiversity value and high threat[20]. Alternatively, species habitat suitability maps can now be published at high resolutions (down to 10 m grid-cell size). Such maps, showing locations that have a high probability of containing

Table 2 | Threats to species related to sharing of biodiversity data

Threat	Examples of the scale of the threat	Direct evidence of data publication leading to population impacts
Collection for international wildlife markets	Mortality and associated population declines due to illegal trafficking of ivory (~5,370 dead elephants) between 1996 and 2008[59]	Published data (digital archives, scientific journals, social media) led to poachers collecting thousands of South African succulents to sell to European plant enthusiasts[a]
Collection for local traditional medicine	Population declines of African pangolins due to harvesting scales/bones for spiritual healing[60]	No published literature found directly linking data publication to impacts
Collection for pet trade	Pet market trade in CITES-listed turtle species in China[61]; illegal import of >20 million live reptiles to the European Union between 2004 and 2014[51]	Local extinctions of newly discovered restricted-range reptiles of value in pet markets following formal publications of locality information[62]
Resource use such as harvesting animals or eggs for food	Declines in marine stocks (for example, sea cucumbers, tuna) due to overfishing[63]; reduced nesting and colony abandonment of Andes flamingos and Peruvian heronries due to egg harvest[37]	No published literature found directly linking data publication to impacts. Anecdotal records of resource declines (for example, in Coral Triangle spawning locations).
Recreational hunting	Areas with highest trophy hunting levels have highest population declines of Tanzanian lions and leopards[64]	Reduced population survival and group cohesion of lions in Africa due to corrupt tourist hunting practices tracking known individuals/prides[34,35]
Cultural beliefs about negative impacts of species	Illegal killing and disturbance of native flying foxes in Australia perceived as a nuisance due to noise, smell and droppings at roost sites[b]	Research data tracking endangered white sharks used by Western Australian government agency to kill sharks in attempt to reduce shark–human interactions[65]
Recreational disturbance for wildlife watching	Habitat degradation and disturbance of animal behaviours by birdwatchers in popular birdwatching locations[36]	Data shared on breeding location of painted snipe pair resulted in nest being abandoned due to disturbance by wildlife enthusiasts[c]
Killing/habitat destruction to prevent conservation	Killing of mountain gorillas in the Democratic Republic of the Congo allegedly by rebels, soldiers, corrupt officials to discredit conservationists and facilitate access to World Heritage region for illegal charcoal trade[66,67]	'Panic clearing' by Australian landholders of vegetation needing protective mechanisms after government report[68] mapped occurrence of rare/threatened communities[d]

[a]http://e360.yale.edu/features/unnatural-surveillance-how-online-data-is-putting-species-at-risk; [b]https://www.animalecologylab.org/not-in-my-backyard.html; [c]http://www.shanghaibirding.com/2017/08/24/painted-snipe-ethics; [d]https://theconversation.com/australia-is-a-global-top-ten-deforester-and-queensland-is-leading-the-way-87259. Making localities of some species available to the public can increase one or more of three major threats to species: poaching (driven by illegal wildlife trade); hunting/collecting for food, medicine and trophies; and habitat disturbance and degradation. Effects on populations of sensitive species range from disturbance of regular behaviours and reduced reproductive rates to reduced population survival and possible local or global extinctions, but there is little direct evidence linking data publication to these impacts. Literature review conducted by searching Web of Science (http://wok.mimas.ac.uk/; accessed 20 August 2017) for keywords 'species' AND 'data publication' or 'data sharing' or 'publication of data' AND 'threat' or 'poach*' or 'wildlife trade' or 'recreational hunting' or 'trophy hunting'.

the species, often include (1) locations where the species occurs and this is known; (2) locations where the species occurs but this is not known; and (3) locations where the species does not occur but can colonize or be translocated if habitat quality is maintained. It is not possible to distinguish between the last two categories a priori so they are typically represented as a combined mapped area (https://mol.org/species/map/). As habitat suitability maps are derived from actual species records they are only meaningful and useful if they are produced using precise rather than denatured locations. Hence it is essential that the experts generating these maps have access to full details of the sightings.

Benefits of publishing biodiversity data

Here we define data publishing as the release of primary biodiversity data (defined earlier), or products based on these that link a taxon to a location at a given time, to public databases for use by others. In addition to direct conservation benefits, publishing biodiversity data has multiple benefits for researchers and society including research verification, public engagement, stimulation of new/collaborative research and informing non-researchers about key ecological or conservation issues[21–25] (Table 1).

For species affected primarily by threats such as climate change and habitat loss, if greater availability of biodiversity data enabled more efficient and cost-effective management decisions, the benefits of revealing population locations may outweigh the overall risk of increasing human exploitation of locations[26]. For instance, habitat loss due to forestry and farming is the most frequent threat to global terrestrial biodiversity[27]. Rare species with poorly known distributions are especially likely to have declined from habitat loss, but new populations are often found in unexpected parts of their former ranges[28,29]. Any known location data are crucial to protect the remaining habitats of such species through activities such as building accurate species habitat suitability models[30], which can be incorporated into conservation planning and management. Accurate species distribution models built on fine-resolution location data could result in more effective conservation measures because they can lead to investment in conservation at locations where species occur but have not been sighted and locations where species do not occur but can be colonized or translocated. Sharing data is particularly helpful for data-deficient species that often slip through the net of regulatory mechanisms due to poor information on where they are and what threatens them[31]. Ignoring these species in conservation plans risks failing to preserve important locations as well as diversity in ecological traits and evolutionary features of biodiversity[32].

Withholding data and records can lead to perverse outcomes for species requiring management to ensure their persistence.

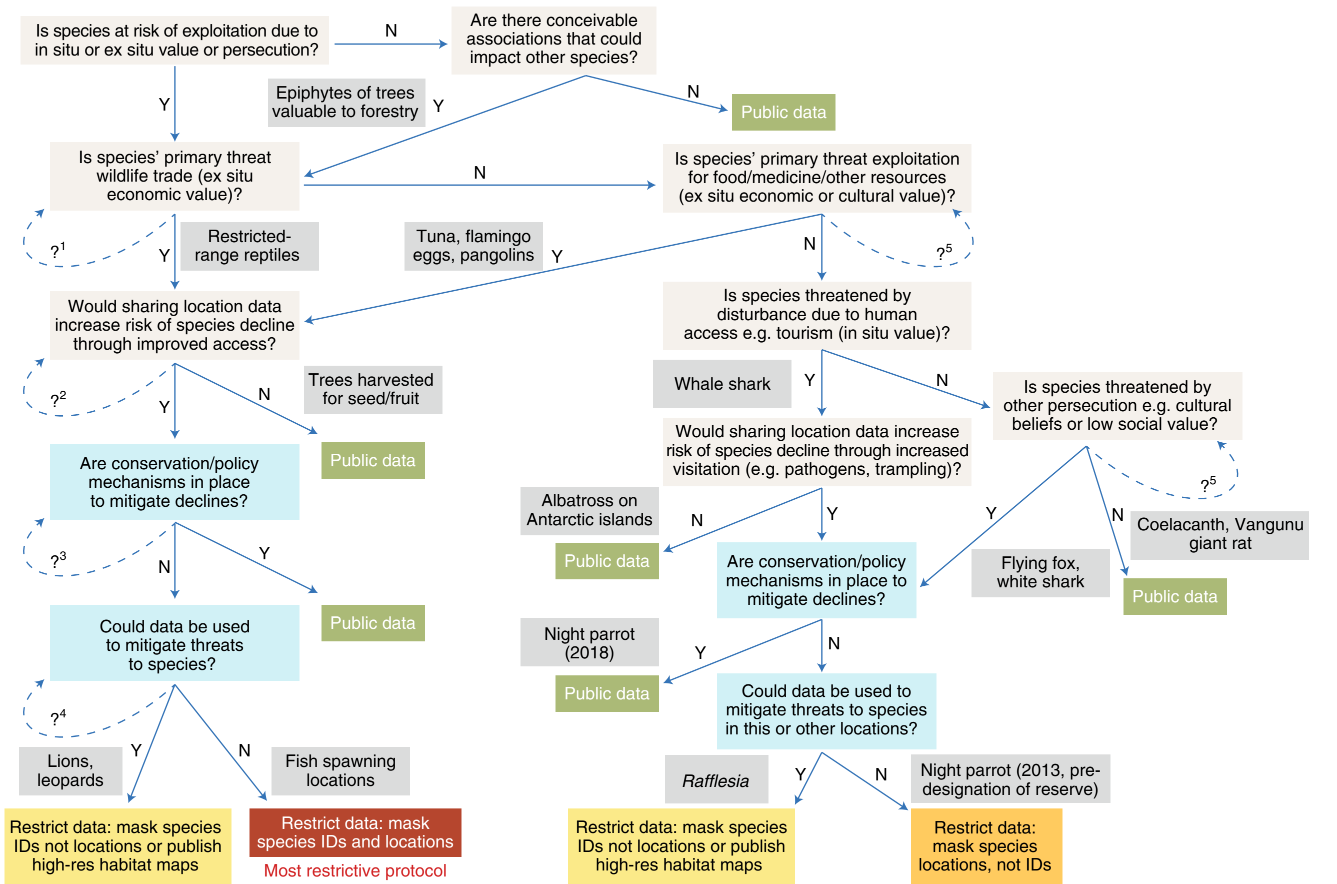

Fig. 1 | Decision tree for publishing biodiversity data from monitoring and surveying. Green, yellow, orange and red boxes indicate data publishing decisions in order of data restrictiveness. Boxes shaded blue indicate considerations of conservation benefits and actions to mitigate impacts resulting from data sharing. Grey boxes indicate examples from the text and Tables 2 and 3. IDs, identities (that is, species' scientific and/or common names). Question marks suggest how to inform particular steps in the tree: ?[1], follow tree according to associated species; ?[1], consult CITES; ?[2], consult IUCN Red List, recovery plans, national/state threat assessments; ?[3], consult global accessibility maps, local people, government threatened species officers in jurisdiction; ?[4], consult conservation evidence and scientific literature; ?[5], consult IUCN Red List, conservation evidence and scientific literature, government threatened species officers in jurisdiction. Y, yes; N, no. For details of terms, see Box 1.

For example, where new locations for threatened species remain undiscovered or are destroyed unknowingly in land development, or there is a false impression of range restriction or small population size. If the objective of government conservation agencies, NGO ecologists, scientists and land managers is to minimize the risk of species extinction (Table 1), then sharing data could help indirectly, by improving information on a species' population size or distribution and enabling a more accurate assessment of threat status, or directly, through enabling increased conservation action in known locations. Additionally, agencies that need occurrence data to manage or assess populations may waste limited resources funding redundant data collection.

Risks of publishing biodiversity data

Despite recent data sharing initiatives and regulations (Supplementary Table 1), there is evidence that different types of data collectors have varying perceptions of how sharing data could undermine their own objectives (Table 1). Moreover, there is no doubt that poaching of species highly valued for traditional medicine and recreational hunting has caused species population declines and even extinctions (for example, the Javan rhino *Rhinoceros sondaicus*[33–35]; Table 2). In addition to documented population declines, human access to habitats has caused individual mortality, changes in wildlife behaviour, reduced reproductive rates and habitat disturbance or loss that affect species' ability to persist in their environment[36,37]. Individual mortality has a greater impact on rare than common species and can cause feedbacks that eventually lead to population declines. Much of the evidence for data publication leading to species declines is anecdotal, with few instances of a direct link between a decline in a population after data on its location being published (Table 2).

Many perceived risks of publishing biodiversity data stem from cultural, social or economic objectives rather than conservation objectives (Table 1). For example, many fishers do not share fishing location data because of concerns their data may be used against them to prosecute for violations or lead to fishing restrictions. Many resource managers view their knowledge as private intellectual property and feel that sharing it with others may put them at an economic and social disadvantage[26]. To achieve a goal of maximizing research output[38], a research scientist might be concerned about the extra time and cost required to share reproducible data, which could instead be used to publish more papers or write more grants.

A balanced decision tree for sharing biodiversity data

From risks to opportunities to conserve species. A sole focus on the risk to a species fails to consider situations in which the benefits of sharing data outweigh the benefits of not allowing access to biodiversity data. The context of any decision about data publication should not miss opportunities to conserve species, and needs to consider public and private costs such as those from redundant

Box 1 | Glossary of terms for the decision tree for publishing biodiversity data from monitoring and surveying

Access. Whether an individual/population can be physically reached by humans; restrictive access can be either due to natural processes/factors (for example, geological = species located on a sea cliff), or human intervention (for example, protection = species in a strictly controlled site).

Associated species. Non-target species that may be dependent on the target species, for example, due to physical/structural relationships (for example, liana on a tropical tree), or nutritional interactions (for example, obligate nectarivore or pollen producer, or obligate seed disperser such as flying fox).

Conservation/policy mechanisms. Actions that aim to improve species/population persistence by mitigating threats causing population declines.

Data repository or archive. A permanent collection of data sets with accompanying metadata usually stored in an online cloud service so that a variety of users can readily access, understand, download and use data[78].

Ex situ economic value. Individuals or populations have monetary worth outside the natural/native site, for example, exploitation through harvesting of individuals for food, medicine or wildlife trade.

Impact on other species. For example, ground-level species may be (inadvertently) trampled/damaged by people walking to access the target species, or arboreal species by people climbing trees to view another target species.

Impact on population. Direct (removal), physical (damage to individuals), or environmental (habitat degradation/reduction) impact, such that population viability is compromised.

In situ value. Individuals/populations have monetary or cultural worth in natural/native site, for example, for ecotourism.

In situ disturbance. Impacts on individuals/populations in natural/native site due to human activity, for example, invasive species, tourism, habitat degradation.

Public data. Information on species locations, abundance, ecology published with no restrictions so that it is openly available to the public (that is, primary data with no denaturing or masking).

Restrict data. Information on species locations, abundance, ecology published only after being masked, generalized or converted to habitat suitability model. Some masked species and location data are likely to be stored offline but models may be publicly available.

surveying effort or the loss of a species. As such, we propose that scientists follow a decision tree that considers the benefits of sharing biodiversity data (Fig. 1 and Box 1), which include highlighting species and places of conservation concern (Table 1). In our decision tree, we assess these kinds of benefit against possible risks of sharing data, such as increased pressure on populations (Table 2). Importantly, our protocol considers all relevant threats to the species, and whether conservation mechanisms are either already in place or could be put in place to mitigate or avoid these. A balanced and transparent evaluation of how, not whether, to share biodiversity data requires owners to clarify the risks and likely impacts to a species from data publication, and at the same time help place this information in a decision-making framework that considers actions to reduce risks of harm to species.

Risk management for species at threat of exploitation. Following a risk assessment approach[39] to sharing biodiversity data, we agree with other discussions on data sharing[16] that it is first necessary to identify the risk of published locality data enabling (or increasing) access to a species based on how valuable and accessible it is to collectors, poachers, recreational visitors, or other people with interest in the species (Supplementary Fig. 1). This will enable those considering publishing spatial biodiversity data to assess the likely harm to the species or population if visitors disturb or exploit it at published localities.

Our protocol accounts for various kinds of risk to species from data publication that have been identified in existing ethical data publication guidelines (Supplementary Table 1). The main risks are increased exploitation for trade or resource use (ex situ threats), or disturbance/destruction of habitat due to human access to localities (in situ impacts). High ex situ value species are those exploited by the wildlife trade or for resources such as food or timber (see the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) Appendix 1 or 2 species; https://cites.org/eng). Well-known examples include the African white and black rhinoceros, all elephant species and many fish. Our decision tree also accounts for the fact that risks to some species might be mitigated by conservation measures, such as restricting access to important sites through regulations or physical barriers (for example, fencing off reserves) — actions that might enable public sharing of data. For species where it is not feasible to restrict access, data publication protocols that mask certain characteristics of the data might be used to protect the species identification or location by the public (Supplementary Table 1), although this might restrict the ability of conservation planners and managers to use the data. We suggest in our decision tree that building a high-resolution habitat model with the data would be a sensible way to publish the data while ensuring the exact locations of individuals were masked (Fig. 1). The full data could be stored securely and granted after request and assessment of motivations, with data protected by a data sharing agreement. For example, a government conservation agency could collate all threatened species occurrence data from researchers licensed to conduct studies on a species, and build a high-resolution map to provide information to the public about the habitat requirements and distribution of species to engage people while not providing site-specific occurrence data that would be available under licence for researchers. For some species, the risk is so high that both measures (masking locations and restricting access) should be enacted while ensuring that monitoring is undertaken to track changes in the species and their threats — this equates to the strictest protocol in our decision tree (Fig. 1). One example is fisheries spawning locations, which are almost impossible to restrict access to when in international waters, but have high value and a history of overharvesting (example 2 in Table 3).

For many high-risk species, releasing public data on occurrence might increase the risk of species decline if locations were previously unknown. If no conservation measures are in place to avoid these declines, our tree suggests that data be desensitized to mask the identities of species but not locations (Fig. 1). This would mean the public can still learn that a precise location has conservation value but would not have specific information on which threatened species occur there, reducing the incentive to visit the location. In some cases, however, sharing location data is unlikely to increase the risk of decline, as population information is already in the public domain, or there is poor access to populations. For these populations, we recommend publishing data without restrictions (Fig. 1),

Table 3 | Examples of how decision tree can be used to support decisions to share data (see Supplementary Table 2 for details)

Species and location	Risks of data sharing	Conservation mechanisms and benefits of data publication	Decision tree action and reasons
1. Night parrot *Pezoporus occidentalis*, arid zone of Australia	No population of live birds known until 2013. Exact location withheld from the public. To limit risks of poaching or disturbance from birdwatchers, very few conservation professionals, government representatives and researchers had access to location data and calls.	In 2016, Queensland State Government made discovery location an exclusion zone with imprisonment if accessed unlawfully. Bird's call and data on habitat preferences released when conservation mechanisms were in place. At least three new populations discovered since by people using recorded call[69,70].	Species was at risk of exploitation due to in situ activities (disturbance due to human activities) in 2013. Decision: Mask species locations, not identities. In 2018, protection mechanisms are in place to mitigate declines. Decision: public data.
2. Fish spawning aggregations, Pacific Ocean	Species that spawn seasonally in large groups are easy for fishers to exploit. Over-exploitation of reef fish spawning aggregations led to collapse of several species following discovery of new aggregation locations, with 20 species of grouper at risk of extinction if current overfishing trends continue.	Controlled release of spawning data to researchers by the NGO Science and Conservation of Fish Aggregations through data sharing confidentiality agreements enabled decision-makers to target spatial conservation planning to protect species such as threatened grouper (Serranidae), snapper (Lutjanidae) and emperor fish (Lethrinidae)[71–75].	Species are at risk of ex situ exploitation for food and fish trade. Conservation/policy mechanisms to mitigate declines not in place at all sites. Location data useful to inform management. Data can be requested by researchers and used under protocols that restrict publishing of maps identifying spawning locations. Decision: restrict data: mask species identities and locations.
3. Vangunu giant rat *Uromys vika*, Solomon Islands	Recently discovered mammal known from a single specimen on Vangunu Island[76]. Location is remote and difficult to access, requiring permission from landowners and tribal chiefs. Forest inhabited by *U. vika* is threatened by logging and not secured under any formal conservation land.	Local community attempting to protect habitat as the Zaira Community Resource Management Area (ZCRMA), which also supports a vulnerable bat *Pteralopex taki* and nesting leatherback turtles *Dermochelys coriacea*. It is hoped that increased recognition will attract further support for ZCRMA and thus protect *U. vika*.	Species has no ex situ economic value and not threatened by disturbance from increased visitation due to remoteness of location. Publication of holotype location data will increase recognition of forests' conservation value and may help conservation efforts to set up protected areas. Decision: public data.
4. Parasitic corpse flower *Rafflesia*, Malaysia	Populations consist of few individuals, are relatively accessible and threatened by logging, agriculture and disturbance by unsustainable ecotourism. Releasing species occurrence data without conservation measures would increase risk of decline of *Rafflesia* and host plants through increased visitation[77].	Visitors to *Rafflesia* blooming events provide nature ecotourism revenue. Populations occur mostly outside protected areas and there are no proposed protection mechanisms or policies to mitigate declines. Data might be used by conservation planners to propose new protected areas.	Individuals have in situ value for tourism, but could be threatened by increased visitation. Publication of location data is needed to identify sites with conservation value, but human traffic must be managed. Decision: restrict data: mask species identities not locations.

as additional information could benefit species by improving the ability to track changes in a population or discover new populations in other locations based on improved knowledge about habitat preferences.

Some species have higher value in situ than ex situ, with lower or non-existent market value. Species with high in situ value often have high ecotourism value (for example, whale sharks, rare birds), and may be directly impacted by human disturbance and pathogen exposure associated with human movement into and out of their habitat (for example, disruption of bird behaviour through electronic bird song playback). Without appropriate conservation measures such as infrastructure or sensitive guidelines for researchers, threatened or rare species with high in situ value are vulnerable to perturbation by human visitors, and we recommend restricting data in a way that prevents disturbance, for instance through publishing a habitat map instead of raw locations (Fig. 1).

Many species may not be directly impacted by exploitation for trade or tourism but are still at risk of indirect in situ impacts if shared data increase visitation to their localities. For example, the surrounding environment could be negatively affected by vegetation disturbance, soil compaction, or introduction of invasive species, or the species might be impacted through associations with other species. These types of disturbance seem minor compared with direct exploitation but can result in a decline in the condition of the surrounding habitat, which may harm health and alter behaviour. For example, a fungal species may not be vulnerable to threats of increased disturbance or wildlife trade, but the tree species with which it shares a symbiotic relationship could suffer high exploitation rates resulting from harvesting for its timber. In this case, the vulnerability of the tree population in addition to the fungus should be considered when assessing how to publish data on the occurrence of the fungus (Fig. 1).

Maximizing data availability to help conservation. Data sharing among researchers, government agencies, NGOs and citizen science groups will improve our knowledge of population trends and ecology, and our ability to protect species from anthropogenic impacts. Many species such as threatened orchids are vulnerable to in situ human recreational activities through habitat degradation, irrespective of whether collection activities are restricted[40,41]. If conservation measures are not in place, increased visitation resulting from the release of new locations for highly valued recreational species could cause local population declines, and we recommend restricting data to mask either species identities or localities (Fig. 1 and example 4 of Malaysian *Rafflesia* in Table 3), depending on whether data have value for mitigating threats. In many cases, however, conservation measures have been implemented to avoid or mitigate declines (for example, the creation of an exclusion zone to eliminate the chance of visitors disturbing the site; see example 1 of the Australian night parrot *Pezoporus occidentalis* in Table 3). In these cases we recommend making occurrence data public to improve conservation, ecological learning and community engagement (Fig. 1).

When a species' or population's primary threats are neither in situ nor ex situ direct exploitation or disturbance (see example 3 of the Vangunu giant rat *Uromys vika* in Table 3), we recommend making data public, due to either little known risk of increased visitation to the site, or little chance that visitation would affect population viability (Fig. 1). Even when a species has in situ value, the risks of increased visitation might be outweighed by the benefits of publishing data. One example is the Critically Endangered West Indian Ocean coelacanth *Latimeria chalumnae*, an ancient fish thought to have been extinct for 60 million years. In 2000, divers observed coelacanths off South Africa's coast, then tagged several individuals[42]. These rare, slow-growing fish are potentially valuable to collectors, but their deep cave habitats are difficult to access and fisheries bycatch poses a much greater threat to survival than poaching[43]. The location data have been made publicly available, triggering widespread interest among scientists, managers and the public. This publicity helped create new marine protected areas, fisheries management measures and a multinational research programme that has generated more than US$6 million in direct government funding, benefitting many additional species in southern Africa (A. Paterson, South African Institute for Aquatic Biodiversity, personal communication).

In some cases, it may be impossible to decide whether a species has value in the wildlife trade or is vulnerable to visitation disturbance. Until protocols can be updated with new data, we recommend a precautionary approach that restricts data publication, such as that taken in the case of the newly rediscovered, endangered night parrot in Australia (Table 3).

Flexibility to adapt to different contexts and changing information. A major challenge of data publication is the evolving nature of restricted data. Lists of 'sensitive' species are useful for some data publication protocols (see Supplementary Table 1), but these lists need regular updating to account for changes in conservation status, knowledge and threats, and need to be adopted on a national or global scale. The IUCN Red List is partially revised each year, but local and regional information on threats to species is often poorly mapped and ad hoc[44]. Genuine status changes may be rapid and can apply to previously unrestricted species of Least Concern. For example, five of the six most prominent and economically valuable formerly common ash tree species in North America entered the IUCN Red List in 2017 as Critically Endangered, due to huge mortality from an invasive insect, driven by warming climate[45]. Current data restrictions could also be lifted if new conservation actions are implemented, such as the habitat protections for night parrots (see above). Decisions to share data should therefore be updated iteratively and quickly.

As the problem of data sharing is complex, our proposed decision tree is not a one-fits-all solution, and we hope that additional inputs by scientists and other stakeholders will enhance its structure and application to diverse decision contexts. Designation of species-specific data sharing rules will need to be adapted to existing pressures found on national or subnational scales. Users of the decision protocol should also ensure that criteria used to assess existing conservation and policy mechanisms to protect species can detect situations where policy mechanisms or legislation exist but are not implemented. We developed our decision tree based on the objective of maximizing persistence of a species, but the protocol could be adapted to account for additional objectives, such as maximizing public engagement in conservation, or conserving whole ecosystems.

Ensuring data re-use and application to conservation decisions. An essential step to promote data sharing and enhance data re-use is to ensure that users know which data exist and are available. Metadata represent the set of instructions or documentation that describe the content, context, quality, structure and reusability of a data set. In addition to publishing biodiversity data, making public the background metadata is critical, and could be accompanied by a sample of the database to enable potential users to assess if those data are fit for purpose[46]. We present our protocol anticipating that repositories holding biodiversity data will have cybersecurity data administrators managing the security of holdings. Data policies should state repository security so that data submitters can decide whether the repository is trustworthy. As species locality data are found in multiple repositories, we recommend that the appropriate mode of sharing biodiversity data should be a species or population attribute rather than an attribute of a given set of data points specified by data authors. This places greater responsibility on researchers to determine how to share data and the decision tree we have proposed should help this.

Although acquiring the information needed to walk through our decision tree could sometimes be time consuming and difficult for individual researchers to obtain, all the information needed for applying our decision tree will be available to those evaluating species for CITES or for IUCN Red Listing. Hence it would make sense for the application of our decision tree to be integrated into these evaluation processes as well as national and subnational assessments of species threat status and updated regularly.

Combating illegal species exploitation

Human exploitation of species for trade, resources, or nature-based recreation continues even in locations with few or no scientific studies. Increased use of social media means that the opportunity to manage sensitive information is declining even if we want to restrict it[47]. The wide range and varied impacts of threats to species mean that researchers and practitioners have an imperative to understand not only where species occur, but also the spread and intensity of both local and off-site threats to species. Despite government agreements such as CITES, illegal resource take (for example, unreported fishing) and wildlife trade continues, with black market prices ranging from US$2 for a sea turtle in Mexico[48], to US$31,000 for an Australian black cockatoo[49], or US$400,000 for a gorilla[50]. It is important to articulate whether these kinds of threat, driven by ex situ markets, are likely to increase when new localities or ecological information on a population are published. In this way, data can be responsibly and appropriately restricted if threats to a species would increase after publishing new localities, or shared without restriction if new data would not affect species persistence (see Fig. 1).

Sharing of species information is without doubt critical in building biodiversity knowledge and managing the global extinction crisis. So far, almost all data publication decisions made by governments, societies or individuals have focused on the costs of sharing; benefits are never explicitly quantified, making it impossible to extrapolate data restriction decisions to other species, locations or contexts. Our decision protocol for publishing spatial biodiversity data aims to overcome this inefficiency and enables scientists to better decide how (and when) to publish data responsibly in repositories. The challenge is to share data in a way that avoids perverse outcomes for biodiversity when it is used. In many cases, sharing data will have greater conservation (and educational) benefits than restricting it from use by those wishing to use it to increase community engagement or to promote conservation actions. Above all else, being explicit about what those benefits might be, and weighing them against the likely risks of making data public, will ensure that species are not put in greater danger from new data being released into the public domain.

Received: 18 February 2018; Accepted: 18 June 2018;
Published online: 23 July 2018

References

1. Costello, M. J., May, R. M. & Stork, N. E. Can we name Earth's species before they go extinct? *Science* **339**, 413–416 (2013).
2. Boitani, L. et al. What spatial data do we need to develop global mammal conservation strategies?. *Phil. Trans. R. Soc. B* **366**, 2623 (2011).
3. Rondinini, C., Wilson, K. A., Boitani, L., Grantham, H. & Possingham, H. P. Tradeoffs of different types of species occurrence data for use in systematic conservation planning. *Ecol. Lett.* **9**, 1136–1145 (2006).
4. Whittaker, R. J. et al. Conservation biogeography: assessment and prospect. *Divers. Distrib.* **11**, 3–23 (2005).
5. Winter, M. et al. Patterns and biases in climate change research on amphibians and reptiles: a systematic review. *R. Soc. Open Sci.* **3**, 160158 (2016).
6. Feeley, K. J. & Silman, M. R. Keep collecting: accurate species distribution modelling requires more collections than previously thought. *Divers. Distrib.* **17**, 1132–1140 (2011).
7. Joppa, L. N. et al. Filling in biodiversity threat gaps. *Science* **352**, 416–418 (2016).
8. Zeller, D., Froese, R. & Pauly, D. On losing and recovering fisheries and marine science data. *Mar. Policy* **29**, 69–73 (2005).
9. Walsh, J. C., Dicks, L. V. & Sutherland, W. J. The effect of scientific evidence on conservation practitioners' management decisions. *Conserv. Biol.* **29**, 88–98 (2015).
10. Turner, D. J., Smyth, A. K., Walker, C. M. & Lowe, A. J. in *Terrestrial Ecosystem Research Infrastructures* (eds Chabbi, A. & Loescher, H. W.) 341–368 (CRC Press, London, 2017).
11. Sullivan, B. L. et al. eBird: a citizen-based bird observation network in the biological sciences. *Biol. Conserv.* **142**, 2282–2292 (2009).
12. Rausher, M. D., McPeek, M. A., Moore, A. J., Rieseberg, L. & Whitlock, M. C. Data archiving. *Evolution* **64**, 603–604 (2010).
13. Whitlock, M. C., McPeek, M. A., Rausher, M. D., Rieseberg, L. & Moore, A. J. Data archiving. *Am. Nat.* **175**, 145–146 (2010).
14. Fuller, R. A., Lee, J. R. & Watson, J. E. M. Achieving open access to conservation science. *Conserv. Biol.* **28**, 1550–1557 (2014).
15. Ngo, H. et al. First population assessment of two cryptic tiger geckos (*Goniurosaurus*) from northern Vietnam: implications for conservation. *Amphib. Reptile Conserv*. **10**, 34–45 (2016).
16. Lindenmayer, D. & Scheele, B. Do not publish. *Science* **356**, 800–801 (2017).
17. Runge, C. A., Tulloch, A., Hammill, E., Possingham, H. P. & Fuller, R. A. Geographic range size and extinction risk assessment in nomadic species. *Conserv. Biol.* **29**, 865–876 (2015).
18. Herzog, S. K., Maillard, Z. O., Embert, D., Caballero, P. & Quiroga, D. Range size estimates of Bolivian endemic bird species revisited: the importance of environmental data and national expert knowledge. *J. Ornithol.* **153**, 1189–1202 (2012).
19. Chapman, A. D. & Grafton, O. *Guide to Best Practices for Generalising Sensitive Species Occurrence Data* (Global Biodiversity Information Facility, Copenhagen, 2008).
20. Alonso, L. E., Deichmann, J. L., McKenna, S. A., Naskrecki, P. & Richards, S. J. *Still Counting...Biodiversity Exploration for Conservation. The First 20 Years of the Rapid Assessment Program* (Univ. Chicago Press, Chicago, 2011).
21. Reichman, O. J., Jones, M. B. & Schildhauer, M. P. Challenges and opportunities of open data in ecology. *Science* **331**, 703–705 (2011).
22. Wallis, J. C., Rolando, E. & Borgman, C. L. If we share data, will anyone use them? Data sharing and reuse in the long tail of science and technology. *PLoS ONE* **8**, e67332 (2013).
23. Borgman, C. L. The conundrum of sharing research data. *J. Am. Soc. Inf. Sci. Tec.* **63**, 1059–1078 (2012).
24. McKiernan, E. C. et al. The open research value proposition: how sharing data can help researchers succeed. *figshare* https://doi.org/10.6084/m9.figshare.1619902.v2 (2016).
25. Borgman, C. L. et al. Knowledge infrastructures in science: data, diversity, and digital libraries. *Int. J. Digit. Libr.* **16**, 207–227 (2015).
26. Johnson, T. R. & van Densen, W. L. T. Benefits and organization of cooperative research for fisheries management. *ICES J. Mar. Sci.* **64**, 834–840 (2007).
27. Maxwell, S. L., Fuller, R. A., Brooks, T. M. & Watson, J. E. Biodiversity: the ravages of guns, nets and bulldozers. *Nature* **536**, 143–145 (2016).
28. Fisher, D. O. Trajectories from extinction: where are missing mammals rediscovered? *Glob. Ecol. Biogeogr.* **20**, 415–425 (2011).
29. Fisher, D. O. & Blomberg, S. P. Correlates of rediscovery and the detectability of extinction in mammals. *Proc. R. Soc. B* **278**, 1090–1097 (2011).
30. Kissling, W. D. et al. Building essential biodiversity variables (EBVs) of species distribution and abundance at a global scale. *Biol. Rev.* **93**, 600–625 (2018).
31. Proença, V. et al. Global biodiversity monitoring: from data sources to essential biodiversity variables. *Biol. Conserv.* **213**, 256–263 (2017).
32. Trindade-Filho, J., de Carvalho, R. A., Brito, D. & Loyola, R. D. How does the inclusion of Data Deficient species change conservation priorities for amphibians in the Atlantic Forest? *Biodivers. Conserv.* **21**, 2709–2718 (2012).
33. Setiawan, R. et al. Preventing global extinction of the Javan rhino: tsunami risk and future conservation direction. *Conserv. Lett.* **11**, e12366 (2018).
34. Leader-Williams, N., Baldus, R. D. & Smith, R. J. in *Recreational Hunting, Conservation and Rural Livelihoods: Science and Practice* (eds Dickson, B., Hutton, J. & Adams, W. A.) 296–316 (Wiley-Blackwell, Oxford, 2009).
35. Loveridge, A. J., Hemson, G., Davidson, Z. & Macdonald, D. W. in *Biology and Conservation of Wild Felids* (eds Macdonald, D. W. & Loveridge, A. J.) 283–304 (Oxford Univ. Press, Oxford, 2010).
36. Steven, R., Pickering, C. & Guy Castley, J. A review of the impacts of nature based recreation on birds. *J. Environ. Manag.* **92**, 2287–2294 (2011).
37. Gonzalez, J. A. Effects of harvesting of waterbirds and their eggs by native people in the northeastern Peruvian Amazon. *Waterbirds* **22**, 217–224 (1999).
38. Arlettaz, R. et al. From publications to public actions: when conservation biologists bridge the gap between research and implementation. *Bioscience* **60**, 835–842 (2010).
39. Burgman, M. A. *Risks and Decisions for Conservation and Environmental Management* (Cambridge Univ. Press, New York, 2005).
40. Kelly, C. L., Pickering, C. M. & Buckley, R. C. Impacts of tourism on threatened plant taxa and communities in Australia. *Ecol. Manag. Restor.* **4**, 37–44 (2003).
41. Ballantyne, M. & Pickering, C. Ecotourism as a threatening process for wild orchids. *J. Ecotourism* **11**, 34–47 (2012).
42. Hissmann, K. et al. The South African coelacanths - an account of what is known after three submersible expeditions: coelacanth research. *S. Afr. J. Sci.* **102**, 491–500 (2006).
43. Erdmann, M. Lessons learned from the conservation campaign for the Indonesian coelacanth, *Latimeria menadoensis*: coelacanth research. *S. Afr. J. Sci.* **102**, 501–504 (2006).
44. Tulloch, V. J. D. et al. Why do we map threats? Linking threat mapping with actions to make better conservation decisions. *Front. Ecol. Environ.* **13**, 91–99 (2015).
45. *IUCN Red List of Threatened Species*. Version 2017-3 (IUCN, Gland, accessed 5 January 2018); http://www.iucnredlist.org
46. Michener, W. K. Ecological data sharing. *Ecol. Inform.* **29**, 33–44 (2015).
47. Smith, M., Szongott, C., Henne, B. & Von Voigt, G. Big data privacy issues in public social media. In *2012 6th IEEE International Conference on Digital Ecosystems Technologies (DEST)* 1–6 (IEEE, 2012).
48. Mancini, A. & Koch, V. Sea turtle consumption and black market trade in Baja California Sur, Mexico. *Endanger. Species Res.* **7**, 1–10 (2009).
49. Barry, C. Australia's wildlife blackmarket trade. *Australian Geographic* (16 August 2011); http://www.australiangeographic.com.au/topics/wildlife/2011/08/australias-wildlife-blackmarket-trade/
50. Stiles, D., Redmond, I., Cress, D., Nellemann, C. & Formo, R. K. *Stolen Apes: The Illicit Trade in Chimpanzees, Gorillas, Bonobos and Orangutans* A Rapid Response Assessment (United Nations Environment Programme, GRID-Arendal, Arendal, 2013).
51. Auliya, M. et al. Trade in live reptiles, its impact on wild populations, and the role of the European market. *Biol. Conserv.* **204A**, 103–119 (2016).
52. Collins-Kreiner, N., Malkinson, D., Labinger, Z. & Shtainvarz, R. Are birders good for birds? Bird conservation through tourism management in the Hula Valley, Israel. *Tour. Manag.* **38**, 31–42 (2013).
53. Lowe, A. J. et al. Publish openly but responsibly. *Science* **357**, 141–141 (2017).
54. Michener, W. K., Brunt, J. W., Helly, J. J., Kirchner, T. B. & Stafford, S. G. Nongeospatial metadata for the ecological sciences. *Ecol. Appl.* **7**, 330–342 (1997).
55. Piwowar, H. A., Day, R. S. & Fridsma, D. B. Sharing detailed research data is associated with increased citation rate. *PLoS ONE* **2**, e308 (2007).
56. Campbell, E. G. & Bendavid, E. Data-sharing and data-withholding in genetics and the life sciences: results of a national survey of technology transfer officers. *J. Health Care Law Policy* **6**, 241–255 (2003).
57. Tenopir, C. et al. Data sharing by scientists: practices and perceptions. *PLoS ONE* **6**, e21101 (2011).
58. Keniger, L., Gaston, K., Irvine, K. & Fuller, R. What are the benefits of interacting with nature? *Int. J. Env. Res. Pub. Health* **10**, 913 (2013).
59. Rosen, G. E. & Smith, K. F. Summarizing the evidence on the international trade in illegal wildlife. *EcoHealth* **7**, 24–32 (2010).
60. Boakye, M. K., Pietersen, D. W., Kotzé, A., Dalton, D.-L. & Jansen, R. Knowledge and uses of African pangolins as a source of traditional medicine in Ghana. *PLoS ONE* **10**, e0117199 (2015).
61. Gong, S.-P., Chow, A. T., Fong, J. J. & Shi, H.-T. The chelonian trade in the largest pet market in China: scale, scope and impact on turtle conservation. *Oryx* **43**, 213–216 (2009).
62. Stuart, B. L., Rhodin, A. G. J., Grismer, L. L. & Hansel, T. Scientific description can imperil species. *Science* **312**, 1137–1137 (2006).

63. Purcell, S. W. et al. Sea cucumber fisheries: global analysis of stocks, management measures and drivers of overfishing. *Fish Fish.* **14**, 34–59 (2013).
64. Packer, C. et al. Effects of trophy hunting on lion and leopard populations in Tanzania. *Conserv. Biol.* **25**, 142–153 (2011).
65. Meeuwig, J. J., Harcourt, R. G. & Whoriskey, F. G. When science places threatened species at risk. *Conserv. Lett.* **8**, 151–152 (2015).
66. Crawford, A. & Bernstein, J. *MEAs, Conservation and Conflicts: A Case Study of Virunga National Park, DRC* (International Institute for Sustainable Development, Geneva, 2008).
67. Jenkins, M. Who murdered the Virunga gorillas?. *Natl Geogr.* **214**, 34–65 (2008).
68. Sattler, P. & Williams, R. *The Conservation Status of Queensland*'s *Bioregional Ecosystems* (QLD Environment Protection Agency, Brisbane, 1999).
69. Kearney, M. R., Porter, W. P. & Murphy, S. A. An estimate of the water budget for the endangered night parrot of Australia under recent and future climates. *Clim. Change Responses* **3**, 14 (2016).
70. Murphy, S. A., Silcock, J., Murphy, R., Reid, J. & Austin, J. J. Movements and habitat use of the night parrot *Pezoporus occidentalis* in south-western Queensland. *Austral Ecol.* **42**, 858–868 (2017).
71. Sadovy de Mitcheson, Y. et al. Fishing groupers towards extinction: a global assessment of threats and extinction risks in a billion dollar fishery. *Fish Fish.* **14**, 119–136 (2013).
72. Jumin, R. et al. From Marxan to management: ocean zoning with stakeholders for Tun Mustapha Park in Sabah, Malaysia. *Oryx* 1–12 (2017).
73. Adams, V. M., Tulloch, V. J. & Possingham, H. P. *Land-Sea Conservation Assessment for Papua New Guinea: A Report on the Work Undertaken to Fulfil the Terms of the Project Review and Integration of the Terrestrial and Marine Program of Works on Protected Areas* (University of Queensland, and Conservation and Environment Protection Authority, 2016).
74. Government of Papua New Guinea *National Marine Conservation Assessment for Papua New Guinea* (Conservation and Environment Protection Authority, 2015).
75. Hamilton, R., Sadovy de Mitcheson, Y. & Aguilar-Perera, A. in *Reef Fish Spawning Aggregations: Biology, Research and Management* (eds Sadovy de Mitcheson, Y. & Colin, P.) 331–369 (Fish & Fisheries Series Vol. 35, Springer, Dordrecht, 2012).
76. Lavery, T. H. & Judge, H. A new species of giant rat (Muridae, *Uromys*) from Vangunu, Solomon Islands. *J. Mammal.* **98**, 1518–1530 (2017).
77. Barcelona, J. F., Pelser, P. B., Balete, D. S. & Co, L. L. Taxonomy, ecology, and conservation status of Philippine *Rafflesia* (Rafflesiaceae). *Blumea* **54**, 77–93 (2009).
78. Olson, R. J. & McCord, R. A. in *Ecological Data: Design, Management, and Processing* (eds Michener, W. K. & Brunt, J. W.) 117–141 (Blackwell Science, Oxford, 2000).

Acknowledgements
A.I.T.T. was supported by an Australian Research Council Discovery Early Career Researcher Award DE170100599. E.B., G.E., N.P.L. and L.R. were supported by the Australian Government National Environmental Science Programme's Threatened Species Recovery Hub. N.P.L. was partially funded by Bush Heritage Australia. N.B. was supported by an Australian Research Council DECRA DE150101552. TERN (A.K.S.) is supported by the Australian National Collaborative Research Infrastructure Strategy. R. Alcorn (eBird), T. Laity (Australian Government Department of the Environment and Energy), S. Murphy and A. Kutt (Bush Heritage Australia) provided feedback on early drafts. J. Miller and R. Fuller contributed to early discussions.

Author contributions
A.I.T.T. led the development of the risk assessment and decision tree with contributions from all authors to the final decision protocol. All authors provided ideas and critical feedback, and co-wrote the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41559-018-0608-1.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence should be addressed to A.I.T.T.