Valerie Racine's Commentary on "Big Data & Conservation Biology"

Commentary On
Case: Big Data & Conservation Biology

Some of the main goals in conservation biology are to track changes in large-scale ecosystems and to conserve biodiversity. Defining and assessing ‘biodiversity’ presents many epistemological challenges to which many scientists attend (cf. Sarkar 2002; Sarkar *et al.* 2006). Moreover, conservation biologists must collect, maintain, and analyze large sets of data. And, with better technology to track and measure biological and environmental variables, and the ability to share or create open access databases, conservation biology faces emerging ethical issues concerning its reliance on big data.

As with other sciences, the use of big data in conservation biology has led to ethical considerations about how to best balance basic scientific virtues, like the open flow of information and collaborations across borders, with the need to protect participant privacy and to maintain confidentiality in certain contexts (Bowser *et al.* 2014).

In the hypothetical scenario described above, the context is a citizen science project in which amateur birders share records of their observations, which are then curated and annotated by experts to become data made available on an open access platform. Because private citizens are sharing information, the eBird website clearly outlines its privacy policy to inform participants that although no personal contact information is publicly available, be aware that any details of an observation as well as its corresponding location (species, numbers, etc.) are in fact available to all users who are registered with eBird (Cornell Lab of Ornithology 2018). Therefore, it is possible that information about sightings on private property become publicly available. In the scenario, the interactions between Andrei and his neighbor, Anna, illustrate one particular ethical issue that might arise from such circumstances.

The potential costs of reporting on the presence of sensitive or endangered species presents another concern related to confidentiality, which also arises in the hypothetical scenario. The eBird website includes guidelines for reporting on sensitive species (Team eBird 2012).

Conservationists worry that publicizing the explicit coordinates or directions to the locations of sensitive or rare species might encourage more traffic in the area, which may lead to an increase in the risk of human disturbance to a vulnerable species' habitats. There are also potential negative effects that reports of rare birds might have on the quality of the databases to which they are submitted. For example, the reports may lead to the phenomena of "twitching" – "the act of making trips specifically to see previously reported rare birds" (Straka & Turner 2013, 40). Twitching can lead to biased samples of checklists or misleading data on bird abundance in open access databases from citizen science projects (Straka & Turner 2013; Kelling *et al.* 2009).

There are additional concerns about confidentiality and security with respect to publicly available data in conservation biology, sustainability, and environmental sciences (Keeso 2014). For example, poachers may gain access to the locations of endangered species and cause harm. Governments are sometimes hesitant to disclose detailed geographical maps – which might be very useful to scientists in tracking a region's biodiversity – for reasons of national security. And, some corporations and scientists are worried about confidentiality because they view their data as proprietary.

Moreover, new technologies used by conservation biologists in the field to gather data, such as biotelemetry, might require interventions in natural habitats, which raise some ethical concerns, especially in the context of research on endangered species or sensitive ecosystems (Cooke 2008; Jewell 2013). The use of biotelemetry often requires tagging individuals of a species. This generates valuable information that may be useful to inform conservation priorities and meet conservation goals, and the assumption is that such interventions will not harm the welfare of individuals or populations, but the risk of harm is still a possibility. To mitigate these harms, biologists have made efforts to weigh the relative benefits of the research and any costs to individuals and populations. Researchers also investigate the impacts of tagging activities and test tagging techniques to develop better intervention practices (Cooke 2008, 172).

Furthermore, researchers have considered some of the large-scale effects of big data biodiversity projects, such as the global biodiversity information facility (GBIF), on the priorities and practices of ecological sciences (Devictor & Bensaude-Vincent 2016). They argue that the conversion of records and observations into data – what they call the process of datafication – results in the loss of information (e.g. de-contextualization) about particular environments or ecosystems, which in turn transforms the science of ecology from one centered on environmental management to one centered at providing and managing data for environmental management. They refer to this shift of focus as a transformation of ecology into a "technoscience" (Devictor & Bensaude-Vincent 2016, 20). This transformation might have harmful implications if it leads to a situation where scientists feel justified in accumulating data and monitoring global diversity without any concern for consequences occurring at smaller, local scales, or concern about the lack of political action needed to protect local environments or ecosystems (Devictor & Bensaude-Vincent 2016, 19-20).

While an emphasis on the accumulation of big data for conservation biology and environmental science might lead to a neglect of local contexts, some researchers have instead emphasized potential ethical upshots and societal benefits of big data, and data sharing in particular, within these fields. For example, Soranno *et al.* (2015) claim that "the issue of data sharing holds the potential for environmental scientists to align their practice with the discipline's growing interest in issues of social consciousness, the democratization of science, inclusion, and scientific literacy" (Soranno *et al*. 2015, 71). According to these authors, the increasing reliance on public participation in, and sponsorship of, research creates an ethical obligation for scientists to promote and facilitate data sharing.

References

1. Cornell Lab of Ornithology. 2018. "Home: Privacy Policy and Cookie Policy." Accessed 14 May 2021. https://www.birds.cornell.edu/home/privacy.

2. Team eBird. 2012. "Guidelines for Reporting Sensitive Species." Accessed 14 May 2021. https://ebird.org/news/sensitive_species.