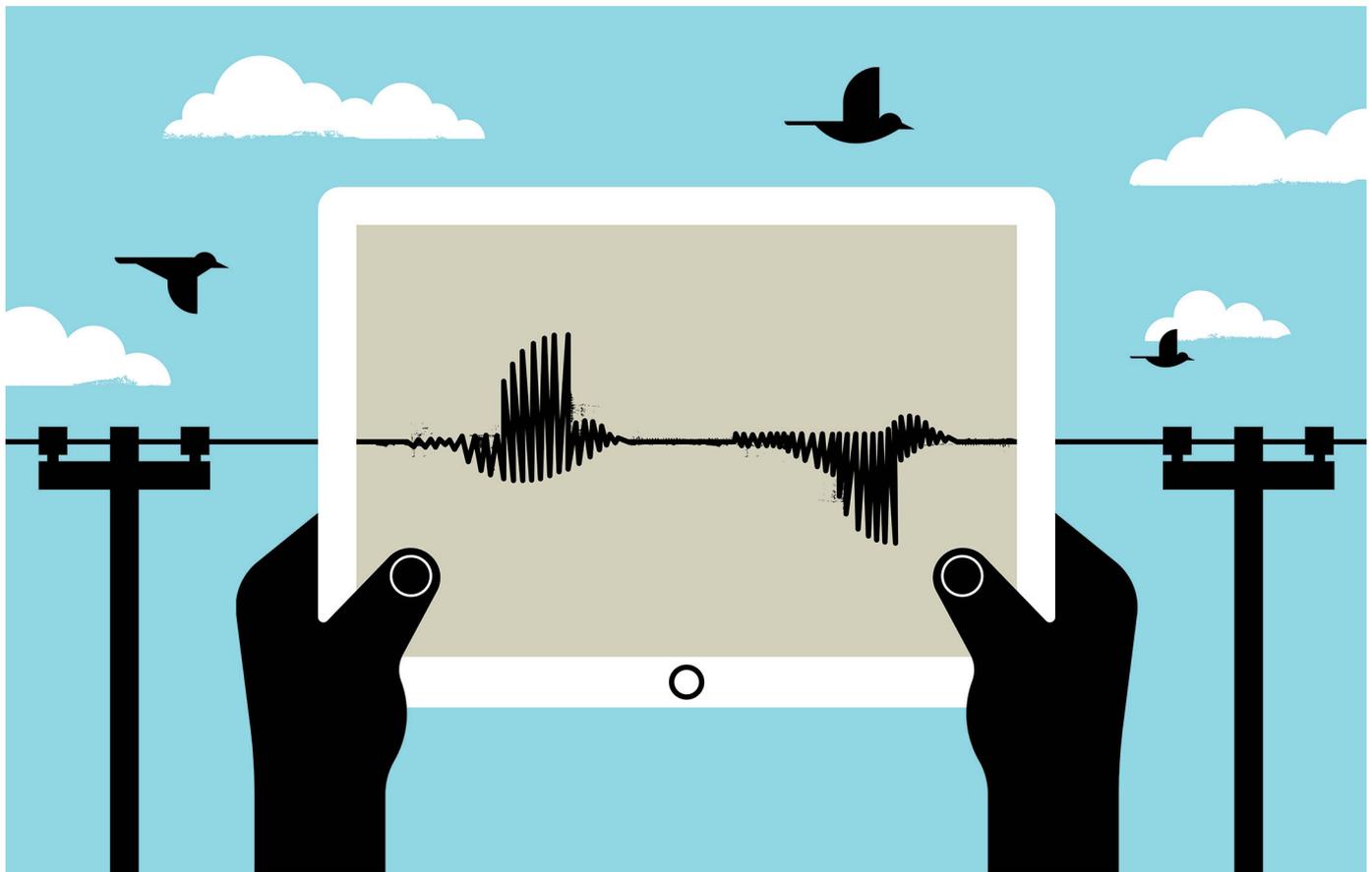


## TOOLBOX

# AI EMPOWERS CONSERVATION BIOLOGY

*Faced with mountains of image and audio data, researchers are turning to artificial intelligence to answer pressing ecological questions.*

ILLUSTRATION BY THE PROJECT TWINS



BY ROBERTA KWOK

When researchers collect audio recordings of birds, they are usually listening for the animals' calls. But conservation biologist Marc Travers is interested in the noise produced when a bird collides with a power line. It sounds, he says, "very much like the laser sound from *Star Wars*".

In 2011, Travers wanted to know how many of these collisions were occurring on the Hawaiian island of Kauai. His team at the University of Hawaii's Kauai Endangered Seabird Recovery Project in Hanapepe was concerned specifically about two species: Newell's shearwaters (*Puffinus newelli*) and Hawaiian petrels

(*Pterodroma sandwichensis*). To investigate, the team went to the recordings.

With some 600 hours of audio collected — a full 25 days' worth — counting the laser blasts manually was impractical. So, Travers sent the audio files (as well as metadata, such as times and locations) to Conservation Metrics, a firm in Santa Cruz, California, that uses artificial intelligence (AI) to assist wildlife monitoring. The company's software was able to detect the collisions automatically and, over the next several years, Travers' team increased its data harvest to about 75,000 hours per field season.

Results suggested that bird deaths as a result of the animals striking power lines numbered in the high hundreds or low thousands, much

higher than expected. "We know that immediate and large-scale action is required," Travers says. His team is working with the utility company to test whether shining lasers between power poles reduces collisions, and it seems to be effective. The researchers are also pushing the company to lower wires in high-risk locations and attach blinking LED devices to lines.

For underfunded conservation scientists, AI provides an attractive alternative to manually processing huge troves of data, such as camera-trap images or audio recordings. A PhD student could "spend months labelling it all by hand before they can get anywhere near answering their hypothesis", says Dan Stowell, a computer scientist at Queen Mary University of ▶

► London. Citizen science is one option to help with the data, but it isn't always the right one: volunteers might work too slowly, and recruiting for projects that involve non-charismatic species can be difficult. AI tools don't experience fatigue-related performance deterioration, as humans do, and they might be better at detecting infrequent or complex patterns.

Scientists need answers to pressing questions, such as whether conservation actions are working. And some problems need near real-time results — for instance, law-enforcement agencies pursuing illegal wildlife traffickers need to determine quickly if an animal for sale on social media is protected. AI could fit the bill.

Despite its reputation for requiring advanced computing skills, AI is now more accessible than ever, thanks to point-and-click tools and dedicated programming libraries. The software isn't as accurate or as sensitive as humans at many conservation research tasks, and the amount of data needed to train an AI algorithm to recognize images and sounds can present hurdles. But early adopters in conservation science are enthusiastic. For Travers, AI enabled a massive boost in monitoring. "It's a huge increase over any other method available," he says.

#### AUTOMATED IDENTIFICATION

Researchers interested in AI can, as Travers did, outsource the work. Conservation Metrics' pricing starts at US\$1–3 per hour of audio, but depends on data volume and project complexity; image-classification pricing is variable. The company also informally takes on for free three to five projects per year that involve interesting conservation questions or technical challenges; researchers can contact the company for consideration.

Scientists can also use browser-based tools. One option is Wildbook, a software framework produced by the non-profit organization Wild Me in Portland, Oregon, and its academic partners. Wildbook uses neural networks and computer-vision algorithms to detect and count animals in images, and to identify individual animals within a species. This information enables more precise estimates of wildlife population sizes, says Tanya Berger-Wolf, Wildbook co-founder and a computer scientist at the University of Illinois at Chicago.

Wildbook works for any species with stripes, spots, wrinkles, fin or ear notches, or other unique physical features. The number of manually annotated images needed to start a project depends on the species, says Jason Holmberg, executive director of Wild Me. The team also uses AI to collect images of specific species from YouTube, and will start trawling through Twitter this year. Researchers and citizen scientists can add images to be identified automatically. So far, scientists have created projects for more than 20 types of animals, including whale sharks (*Rhincodon typus*), manta rays (*Manta birostris* and *M. alfredi*), Iberian lynxes (*Lynx pardinus*) and giraffes (*Giraffa* sp.). New users can join existing Wildbook projects for free; setting

up a project for a new species typically costs \$10,000–20,000, Holmberg says.

Other tools process acoustic data. The Automated Remote Biodiversity Monitoring Network (ARBIMON) is a browser-based tool produced by Sieve Analytics in San Juan, Puerto Rico. Researchers upload their recordings, and the company recommends that they manually identify a couple of hundred clips that contain the species call of interest, and hundreds more that do not. ARBIMON then uses machine learning to classify the remaining data. So far, about 3.4 million 1-minute recordings have

**“The program doesn't have to be 100% accurate. We just have to know exactly how accurate it is.”**

Sweden, has even incorporated AI software into custom hardware. The firm created a device called the Fish Face ID Tunnel to help researchers identify fish subspecies on the basis of photographs taken inside the device. “You could say it's a photo booth for fish,” says Johanna Reimers, the firm's chief executive. Refind and the environmental charity The Nature Conservancy, based in Arlington, Virginia, installed the device on a fishing boat in Indonesia last year, and the charity is analysing the data. Researchers interested in a similar device, or a customized version, should expect to pay around \$50,000, Reimers says.

#### TRAINING DATA

Tech-savvy researchers can take advantage of 'command-line' AI software to answer conservation questions. Most cutting-edge results in deep learning use the open-source machine-learning libraries TensorFlow, developed by Google, and PyTorch, led by Facebook, Stowell says. For the past few years, his team has run contests, called the Bird Audio Detection Challenge, in which participants develop and test algorithms to determine whether bird calls are present in sound clips. Prizes go to the highest-scoring open-source entries; code from some of the other entries is also posted online.

TensorFlow in particular has a large user community, and many code examples are available online, says Scott Loarie, co-director of the nature app iNaturalist, a joint initiative of the National Geographic Society in Washington DC and the California Academy of Sciences in San Francisco.

Conservation biologists can get up to speed with AI through online classes offered by DataCamp, Coursera, Udacity or the NVIDIA Deep Learning Institute. The online guide Machine Learning for Humans (see [go.nature.com/2sbjasb](http://go.nature.com/2sbjasb)) offers resources for beginners, and the podcast *This Week in Machine Learning & AI* (see [go.nature.com/2ts36bv](http://go.nature.com/2ts36bv)) can help

biologists to stay up to date.

Collaborations with AI specialists can also help. Ruth Oliver, an ecologist at Yale University in New Haven, Connecticut, took this approach when analysing about 1,200 hours of audio from the Arctic. She worked with an electrical engineer, who had experience in machine learning, and ran AI algorithms using the software tool MATLAB to estimate when songbirds arrived in the Arctic and how environmental factors affected vocalizations (R. Y. Oliver *et al. Sci. Adv.* 4, eaaq1084; 2018).

Despite its apparent ubiquity, AI isn't an easy solution. Estimates of the amount of training data required for machine learning vary widely, from hundreds to tens of thousands of manually classified examples, depending on the model, study goals and task complexity. But for endangered species, it might be “difficult or impossible” to collect a very large sample, says Mitch Aide, a tropical ecologist at the University of Puerto Rico Río Piedras Campus in San Juan and founder of Sieve Analytics. “In these cases, you do the best you can.”

AI software can also be more error-prone than trained people. Stowell suggests validating results with a small sample and testing the model on data from a different research project or country to ensure that it generalizes as expected.

To assess accuracy in its work, Travers' team performed extensive tests comparing Conservation Metrics' results with on-the-ground observations. The AI software picked up about half of the strikes seen and heard by humans, and an analyst at the firm manually removed false positives. To estimate the actual collision rate, the scientists roughly doubled the number detected by the software. The AI program “doesn't have to be 100% accurate”, Travers says. “We just have to know exactly how accurate it is.”

And then there's the concern that users might blindly accept AI results without understanding how they were generated. To mitigate that, Peter Ersts, a software developer at the American Museum of Natural History's Center for Biodiversity and Conservation in New York City, suggests using open-source tools; biologists could ask a colleague with software-development expertise to review and explain the code.

Ersts and his colleagues hope to build a curated set of labelled wildlife images that researchers can use to test new models. Their open-source program Andenet-Desktop allows users to label species manually and export training data in formats that can be read by frameworks such as TensorFlow and PyTorch. Researchers can use those libraries to create their machine-learning model, load model parameters back into Andenet and automatically annotate remaining data.

Although early results are promising, human input is still needed. “We can't fully replace people yet, and nor should we,” Ersts says. ■

**Roberta Kwok** is a freelance science writer in Kirkland, Washington.