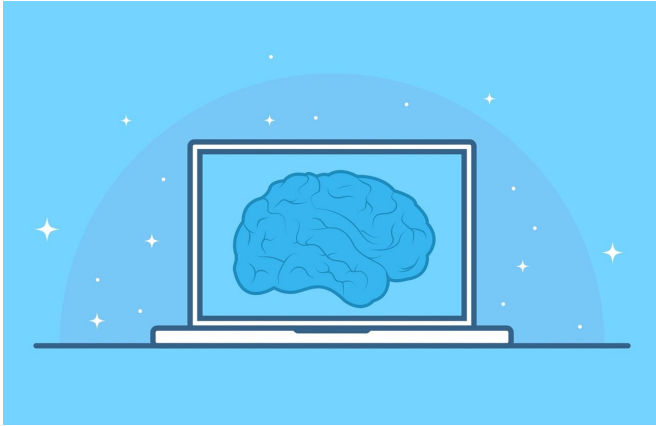


We're getting better at wildlife conservation, AI study of scientific abstracts suggests

19 March 2020



Credit: CC0 Public Domain

Researchers are using a kind of machine learning known as sentiment analysis to assess the successes and failures of wildlife conservation over time. In their study, appearing March 19 in *Patterns*—a new open access data science journal from Cell Press—the researchers assessed the abstracts of more than 4,000 studies of species reintroduction across four decades and found that, generally speaking, we're getting better and better at reintroducing species to the wild. They say that machine learning could be used in this field and others to identify the best techniques and solutions from among the ever-growing volume of scientific research.

"We wanted to learn some lessons from the vast body of [conservation](#) biology literature on reintroduction programs that we could use here in California as we try to put sea otters back into places they haven't roamed for decades," says senior author Kyle Van Houtan, chief scientist at Monterey Bay Aquarium. "But what sat in front of us was millions of words and thousands of manuscripts. We wondered how we could extract data from them that we could actually analyze, and

so we turned to natural language processing."

Natural language processing is a kind of [machine learning](#) that analyzes strings of human language to extract useable information, essentially allowing a computer to read documents like a human. Sentiment analysis, which the researchers used in this paper, looks more specifically at a trained set of words that have been assigned a positive or negative emotional value in order to assess the positivity or negativity of the text overall.

The researchers used the database Web of Science to identify 4,313 species reintroduction studies published from 1987 to 2016 with searchable abstracts. Then they used several "off-the-shelf" [sentiment analysis](#) lexicons—meaning that the words in them had already been assigned a sentiment score based on things like movie and restaurant reviews—to build a model that could give each abstract an overall score. "We didn't have to train the models, so after running them for a few hours we all of a sudden had all these results at our disposal," says Van Houtan. "The scores gave us a trend over time, and we could query the results to see what the sentiment was associated with studies on pandas or on California condors or coral reefs."

The trends they saw suggested greater conservation success. "Over time, there's a lot less uncertainty in the assessment of sentiment in the studies, and we see reintroduction projects become more successful—and that's a big takeaway," he says. "Looking at thousands of studies, it seems like we're getting better at it, and that's encouraging."

"If we are going to maximize our conservation dollars, then we need to be able to quickly assess what works and what doesn't," says study co-author Lucas Joppa, Chief Environmental Officer at Microsoft. "Machine learning, and [natural language](#) processing in particular, has the ability to sift through results and shine a light on [success stories](#)

that others can learn from."

To ensure their results were accurate, the researchers looked at the most common indicators of positive sentiment (and therefore conservation success) in their results and found words like "success," "protect," "growth," "support," "help," and "benefit"; words that indicated negative sentiment were ones like "threaten," "loss," "risk," "threat," "problem," and "kill." These words aligned with what they, as long-time conservation biologists, would typically use to indicate success and failure in their own studies. They also found that trends described by the sentiment analysis for specific reintroduction programs known to be successes or failures (like the reintroduction of the California condor) matched the known outcomes.

The researchers say that off-the-shelf sentiment analysis worked surprisingly well for them, likely because many words used in conservation biology are part of our everyday lexicons and were therefore accurately coded with the appropriate sentiment. In other fields, they think more work would need to be done to develop and train a model that could accurately code the [sentiment](#) of more technical, field-specific language and syntax. Another constraint, they say, is that only a limited number of the papers they sought to analyze were open access, which meant they had to assess abstracts rather than full papers. "We're really just scratching the surface here, but this is definitely a step in the right direction," says Van Houtan.

Still, they do think this is a technique that can and should be applied more broadly in both conservation biology and other fields to make sense of the vast amounts of research that's now being conducted and published. "So much local conservation work goes unnoticed by the global conservation community, and this paper shows how machine learning can help close that information gap," says Joppa.

"Many of these techniques have been in use for over a decade in commercial settings, but we're hoping to translate them into settings like ours to combat climate change or plastic pollution or to promote endangered species conservation," Van Houtan says. "There's a plethora of data that's right

at our fingertips, but it's this sleeping giant because it isn't properly curated or organized, which makes it challenging to analyze. We want to connect people with ideas, capacity, and technical solutions they might not otherwise encounter so we can bring some progress to these seemingly intractable problems."

More information: *Patterns*, Van Houtan et al.: "Sentiment analysis of conservation studies captures successes of species reintroductions" [www.cell.com/patterns/fulltext...](http://www.cell.com/patterns/fulltext...2666-3899(20)30005-2) [DOI: 10.1016/j.patter.2020.100005](https://doi.org/10.1016/j.patter.2020.100005)

Provided by Cell Press

APA citation: We're getting better at wildlife conservation, AI study of scientific abstracts suggests (2020, March 19) retrieved 4 December 2020 from <https://techxplore.com/news/2020-03-wildlife-ai-scientific-abstracts.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.