THIS WEEK

EDITORIALS

WORLD VIEW Opioid epidemic is a sign that the rural United States needs help **p.165**

EXOTIC ATOM Physicists find long-lived variant of unstable potassium **p.167**

NUCLEAR TENSIONS Iran speeds up centrifuge development **p.168**

Take action to stop Amazon fires

The headlines are fading, but the planet's largest rainforest is still on fire. Brazil and the world must halt the destruction before it becomes too late.

Less than a decade ago, Brazil was an environmental leader. Its government had elevated forest conservation and sustainable development to national policy and then, with the help of satellite imagery, it had cracked down on illegal deforestation across the world's largest tropical rainforest. Deforestation in the Amazon plummeted even as agricultural production — the biggest driver of forest loss — increased. Now, that progress is going up in smoke.

Data from Brazil's National Institute for Space Research (INPE) showing a sharp uptick in the number of fires in the Amazon this year triggered headlines around the globe. Landowners use fire to clear forest illegally to make way for crops and cattle grazing, but Brazil's populist president, Jair Bolsonaro, has effectively fanned the flames with his anti-environmentalist agenda since taking office in January. Scientists who live and work in the region were not surprised at what is happening, but INPE's report sparked concerns in world capitals just as leaders of the G7 group of countries with the world's biggest economies gathered for their annual summit in Biarritz, France.

Neither extinguishing the flames nor solving the underlying problem of deforestation will be easy. It doesn't help that Bolsonaro is among those world leaders questioning whether an environmental agenda can deliver long-promised economic benefits. His development-at-any-cost policies hark back to an earlier era in which deforestation was treated as a measure of progress.

He has railed against regulation, cut the budget of Brazil's environmental enforcement agency and advocated mining on lands belonging to Indigenous people. When news of the fires spread, Bolsonaro accused environmental groups of setting blazes to make him look bad. When G7 leaders pledged emergency funding to help put the fires out, he called it colonialism.

MORE EFFORTS NEEDED

The attention of world leaders on the Amazon is welcome, but their response is insufficient to deal with the scale of the crisis. The G7's offer of US$22 million, initially rebuffed by Bolsonaro, seemed rushed. This sum would hardly fight the fires, let alone address the underlying problems. In the words of the former UN climate-secretariat chief Christiana Figueres, it was "a drop in the bucket". On 6 September, at a forest-conservation summit convened by Brazil, seven Amazon countries pledged to work together — but provided few details on what they would actually do.

Paradoxically, there is already a large pot of money dedicated to tropical-forest conservation in Brazil. This is the Amazon Fund, established by Brazil in 2008 to attract international donations for conservation efforts. Since the fund's inception, Norway has invested the lion's share of the almost $1.3-billion total, while Germany has contributed another $68 million and Brazil's national oil company, Petrobras of Rio de Janeiro, nearly $8 million. The funds have been used to pay for everything from research and land-use planning to law enforcement. But these investments were contingent on the government curbing deforestation, and both Germany and Norway have now suspended payments.

This decision is unlikely to change unless there is a shift in the Bolsonaro government's priorities, but European Union countries could have some extra leverage. The EU has negotiated a trade agreement with several South American states, including Brazil. France and Ireland have threatened to refuse to ratify the deal — limiting Brazil's exports of beef and soya to the EU — unless Bolsonaro changes his approach to the Amazon. Brazil's agribusinesses are concerned about these developments. That gives them an opportunity to persuade Bolsonaro to re-engage with Europe over the Amazon if not doing so means that the interests of the country's agricultural producers are on the line.

"Brazil rightly claims sovereignty over its territory, but the forest is a global good."

Fifteen years ago, many people assumed that the Brazilian government had little control or influence over illegal deforestation in the Amazon. We now know that is not true. Between 2004 and 2012, Brazil was able to curb deforestation by more than 80% while almost eliminating industrial-scale land-clearing.

The Amazon rainforest is a reservoir of biodiversity and carbon, which is locked up in trees and soils. Clearing and burning the forest to make way for agriculture destroys the former and sends the latter into the atmosphere, contributing to global warming. Brazil rightly claims sovereignty over its territory, but the forest is a global good, just as the soya beans and beef produced by farmers and ranchers there are global commodities. The responsibility for what happens on Brazil's turf extends well beyond its borders. ■

Free citation data

Most scientists want to judge citation metrics for themselves. That requires data to be accessible.

Whenever scientists are ranked and rewarded by metrics, such as citations, some are tempted to grab a little extra credit where they can. As we report this week, the publisher Elsevier has been investigating cases in which reviewers have repeatedly asked authors of papers to cite the reviewers' own work (page 174).

This is not an isolated incident. Last month, we reported that some 250 highly cited scientists had amassed more than half of their citations from their own work or that of co-authors — much more than

the usual proportion for their field or career stage (see *Nature* **572**, 578–579; 2019).

Such examples should not come as a surprise, because the gaming of measurement systems is well known. In economics it is called Goodhart's law, named after the economist Charles Goodhart, who described the concept. It was refined by the anthropologist Marilyn Strathern, and states that when a measure becomes a target, it ceases to be a good measure.

One obvious answer is for institutions and funders to just stop using citation-based metrics as a proxy for importance or quality when assessing researchers. "Stop the damn bean-counting!" one reader exclaimed in response to an online poll in *Nature* last month, in which we asked what — if anything — needed to be done to curb excessive self-citation. Metrics-based analysis can certainly reveal useful insights about research. But any assessment procedure that rewards scientists according to citation-based metrics alone seems designed to invite game-playing.

It can also be argued that, all things considered, excessive self-citation is a minor problem and therefore doesn't need a particular response. Of the more than 5,000 readers who answered *Nature*'s poll, 10% said nothing needed to be done. "Let active researchers draw their own conclusions about self-citing researchers, and allow reputation to build naturally," one respondent wrote.

However, most poll respondents felt that citation-based indicators are useful, but that they should be deployed in more nuanced and open ways. The most popular responses to the poll were that citation-based indicators should be tweaked to exclude self-citations, or that self-citation rates should be reported alongside other metrics (see 'The numbers game'). On the whole, respondents wanted to be able to judge for themselves when self-citations might be appropriate, and when not; to be able to compare self-citation across fields; and more.

"Free access to citation data will help to illuminate some darker corners."

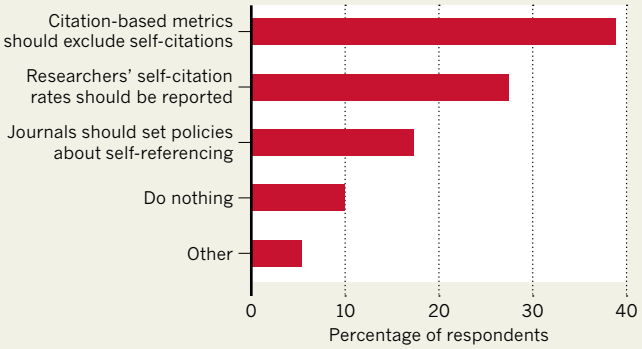
But this is where there is a real problem, because for many papers citation data are locked inside proprietary databases. Since 2000, more and more publishers have been depositing information about research-paper references with an organization called Crossref, the non-profit agency that registers digital object identifiers (DOIs), the strings of characters that identify papers on the web. But not all publishers allow their reference lists to be made open for anyone to download and analyse — only 59% of the almost 48 million articles deposited with Crossref currently have open references.

There is, however, a solution. Two years ago, the Initiative for Open Citations (I4OC) was established for the purpose of promoting open scholarly citation data. As of 1 September, more than 1,000 publishers were members, including Sage Publishing, Taylor and Francis, Wiley and Springer Nature — which joined last year. Publishers still to join I4OC include the American Chemical Society, Elsevier — the largest not to do so — and the IEEE.

Last January, I4OC co-founder David Shotton at the Oxford e-Research Centre, University of Oxford, UK, urged all research publishers to join the initiative (see *Nature* **553**, 129; 2018). They should. Excessive self-citation cannot be eliminated, but free access to citation data for everyone — researchers and non-researchers — will help to illuminate some darker corners. Without more journals coming on board, these necessary efforts to analyse self-citation data will remain incomplete. ■

THE NUMBERS GAME

A *Nature* poll asked what (if anything) should be done to curb excessive self-citation. Respondents said that citation-based indicators are useful, but should be deployed in more nuanced and open ways.



Of 5,575 respondents, 2,183 said citation metrics such as the h-index should exclude self-citations; 1,541 said researchers' self-citation rates should be reported; 968 said journals should set policies about appropriate levels of self-referencing; 565 said to do nothing and 318 chose 'other'.

Material concerns

Materials science embraced machine learning, but researchers must watch for biased data.

Like most research fields, materials science has embraced 'big data', including machine-learning models and techniques. These are being used to predict new materials and properties, and devise routes to existing drugs and chemicals.

But machine learning requires training data, such as those on reagents, conditions and starting materials. These are usually gleaned from the literature, and are human-generated. The choice of reagents that researchers use could come, for example, from experience or from previously published work. It might be based on a recommendation passed from supervisor to graduate student, or simply on how easy reagents are to find or buy. But that subjectivity becomes a potential problem for the accuracy of machine-learning models, as research published this week in *Nature* shows.

Joshua Schrier at Fordham University in New York City, Alexander Norquist and Sorelle Friedler at Haverford College in Pennsylvania and their colleagues looked at materials called amine-templated vanadium borates. These were chosen because success and failure are easily defined in their synthesis — simply by whether or not crystals form. The researchers compiled a data set of several hundred synthetic conditions that are used to make vanadium borates. They then trained a machine-learning model on this data set to predict the success or failure of reactions. The team found that a model trained on a human-generated data set was less successful in predicting the success or failure of a reaction than one trained on a data set with randomly generated reaction conditions (X. Jia *et al. Nature* **573**, 251–255; 2019).

In some sense, this should be no surprise. It is now well known that when machine-learning techniques are used to pick out patterns in aggregated data, biases in those data can be amplified. For example, facial-recognition algorithms trained mostly on white faces are less able to distinguish between the faces of people of other ethnicities, thereby introducing bias that could lead to entrenched inequality.

Does the existence of bias matter to chemistry and materials science? When the goal of a research project is to find new materials, it could be argued that it's irrelevant which reagents are used as long as they work.

But there are potential drawbacks to relying on 'tried and trusted' methods. A prevalence of favourite protocols — even an unintentional one — in a training data set could hinder the success of machine-learning models that are used to predict materials. Or, as this study reveals, more efficient ways to make existing ones.

No one would argue that the consequences of biased chemical data are as serious as those of biases in facial-recognition software, but they share a similar origin. Researchers should be alert to the potential for bias in their chemical data sets, before it gets baked into a machine. ■