



NASA/JPL-CALTECH

## EAR TO THE GROUND

NASA's Mars InSight lander will gather data on seismic activity to help scientists better understand the red planet's mysterious interior.

The 1.8-metre robotic arm will place a seismometer and heat probe onto Mars's surface.

Two solar panels will supply power to the lander and its instruments.

The lander's seismometer will listen for tremors known as marsquakes.

A heat probe will dig down 5 metres to measure temperature change over depth and time.

result of jostling by the wind or other sources. On the third day after landing, project scientists will switch on an instrument to track changes in the magnetic field, which will help them to identify sources of noise that aren't quakes, says Catherine Johnson, a geophysicist at the University of British Columbia in Vancouver, Canada.

InSight won't deploy its German-built heat-flow probe until January. Over the course of several weeks, the instrument will drill five metres into the Martian surface, deeper than anything achieved before. Scientists will track changes in temperature as small as a few hundredths of a degree. That will tell them how much heat is leaving Mars, and how many heat-producing radioactive elements are packed inside it.

InSight is meant to work for a little more than one Martian year, equivalent to almost two Earth years. It should measure 50–100 marsquakes during that period, says Banerdt. The longer it survives, the more it will be able to detect — and the more researchers will be able to deduce about Mars's internal structure. ■

dome-shaped wind shield over it. The whole process is expected to take several days.

The seismometer includes three ground-motion sensors nested inside a vacuum, and

its sensitivity allows it to detect movement as small as the width of an atom. The big challenge will be determining which movements are caused by marsquakes and which are the

## PUBLISHING

# The age of AI peer reviews

*Automated software can help review papers, but the decision-making stays with humans.*

BY DOUGLAS HEAVEN

Most researchers have good reason to grumble about peer review: it is time-consuming and error-prone, and the workload is unevenly spread, with just 20% of scientists taking on most reviews.

Now peer review by artificial intelligence (AI) is promising to improve the process, boost the quality of published papers — and save reviewers time. A handful of academic publishers are piloting AI tools to do anything from selecting reviewers to checking statistics and summarizing a paper's findings.

In June, software called StatReviewer, which checks that statistics and methods in manuscripts are sound, was adopted by Aries Systems, a peer-review management system owned by Amsterdam-based publishing giant Elsevier. And ScholarOne, a peer-review platform used by many journals, is teaming up with UNSILO of Aarhus, Denmark, which uses natural language processing and machine learning to analyse manuscripts.

UNSILO uses semantic analysis of the manuscript text to extract what it identifies as the main statements. This gives a better overview of a paper than the keywords typically submitted by authors, says Neil Christensen, sales director at UNSILO. "We find the important

phrases in what they have actually written," he says, "instead of just taking what they've come up with five minutes before submission."

UNSILO identifies which of these key phrases are most likely to be claims or findings, giving editors an at-a-glance summary of the results. It also highlights whether the claims are similar to those from previous papers, which could be used to detect plagiarism or simply to place the manuscript in context with related work in the wider literature. "The tool's not making a decision," says Christensen. "It's just saying: 'Here are some things that stand out when comparing this manuscript with everything that's been published before. You be the judge.'"

"It doesn't replace editorial judgement but, by God, it makes it easier," says David Worlock, a UK-based publishing consultant who saw the UNSILO demonstration at the Frankfurt Book Fair in Germany last month.

Worlock notes that there are several similar tools emerging. He is on the board of Wizdom.ai in London, a start-up owned by publishers Taylor & Francis, which is developing software that can mine paper databases and extract connections between different disciplines and

concepts. He says that this kind of tool will be useful beyond peer review, for tasks such as writing grant applications or literature reviews.

Many platforms, including ScholarOne, already have automatic plagiarism checkers. And services including Penelope.ai examine whether the references and the structure of a manuscript meet a journal's requirements. Some can flag up issues with the quality of a study, too. The tool statcheck, developed by Michèle Nuijten, a methodologist at Tilburg University in the Netherlands, and her colleagues, assesses the consistency of authors' statistics reporting, focusing on *P* values. The journal *Psychological Science* runs all its papers through the tool, and Nuijten says that other publishers are keen to integrate it into their review processes.

When Nuijten's team analysed papers published in psychology journals, they found that roughly 50% contained at least one statistical inconsistency (M. B. Nuijten *et al. Behav. Res. Meth.* **48**, 1205–1226; 2016). In one in eight papers, the error was serious enough that it could have changed the statistical significance of a published result. "That's worrisome," she says. She's not surprised that reviewers miss such mistakes, however. "Not everyone has time to go over all the numbers. You focus on the main findings or the general story."

For now, statcheck is limited to analysing ►

**"It doesn't replace editorial judgement but, by God, it makes it easier."**



Automation of standardized tasks could take the slog out of peer review.

► manuscripts that use the American Psychological Association's reporting style for statistics. By contrast, the creators of StatReviewer, Timothy Houle at Wake Forest University School of Medicine in North Carolina and Chadwick DeVoss, chief executive of tech start-up NEX7 in Madison, Wisconsin, say that their tool can assess statistics in standard formats and presentation styles from multiple fields. To do this, it checks that papers correctly include things such as sample sizes, information about blinding of experiments and baseline data.

#### DETECTING FRAUD MARKERS

StatReviewer can also identify markers of fraudulent behaviour, says DeVoss. "Things like, did they game some statistical rules, or

did they flat-out make up data? If the risk is higher than what the journal is used to seeing, they can look into the details." DeVoss says that StatReviewer is being tested by dozens of publishers. A 2017 trial with the open-access publisher BioMed Central in London was inconclusive because the tool did not analyse enough manuscripts, but did nonetheless provide some insights. BioMed Central is now planning a follow-up.

StatReviewer did catch things that human reviewers missed, says Amy Bourke-Waite, communications director for open research at Springer Nature, which owns BioMed Central and publishes *Nature* (*Nature's* news team is editorially independent of Springer Nature). For example, it was good at catching papers that did

not meet required standards, such as following CONSORT, a manuscript format used by many publishers. Bourke-Waite adds that authors who took part said that they were as happy responding to StatReviewer reports as they were to the human reviewer's. Occasionally, she says, StatReviewer got things wrong — but sometimes its slip-ups drew authors' attention to unclear reporting in their manuscripts.

Even if the trials prove successful, DeVoss expects that only some journals will want to pay to have all their manuscripts scanned. So he and his colleagues are targeting authors, too, hoping that they will use the tool to check their manuscripts before submission.

There are potential pitfalls to AI in peer review in general. One concern is that machine-learning tools trained on previously published papers could reinforce existing biases in peer review. "If you build a decision-making system based on the articles which your journal has accepted in the past, it will have in-built biases," says Worlock. And if an algorithm provides a single overall score after evaluating a paper, as StatReviewer does, there might be a temptation for editors to cut corners and simply rely on that score in deciding to reject a paper, says DeVoss.

Algorithms are not yet smart enough to allow an editor to accept or reject a paper solely on the basis of the information they extract, says Andrew Preston, co-founder of Publons, a Wellington-based start-up acquired by Clarivate Analytics in Philadelphia, Pennsylvania, that tracks peer review and is using machine learning to develop a tool to recommend reviewers. "These tools can make sure a manuscript is up to scratch, but in no way are they replacing what a reviewer would do in terms of evaluation."

Nuijten agrees: "The algorithms are going to need some time to perfect, but it makes sense to automate a lot of things, because a lot of things in peer review are standard." ■

MARY EVANS/CLASSICSTOCK/H. ARMSTRONG ROBERTS

#### COMMUNITY

# Can conference shed reputation for hosting sexist behaviour?

*AI meeting wants to become more inclusive, but survey suggests it has a long way to go.*

BY HOLLY ELSE

Hordes of artificial-intelligence researchers will descend this weekend on one of the field's hottest tickets: the Neural Information Processing Systems conference in Montreal, Canada. But although attendees at this annual event

will hear talks on cutting-edge ideas in computer science, another issue will also be front and centre: whether the conference can provide a welcoming environment for women as the field of artificial intelligence (AI) grapples with a culture of harassment and discrimination.

The concerns were thrown into stark relief

earlier this month with the release of a survey of 2,375 people — most of whom had either attended the meeting or submitted papers for consideration in previous years.

Respondents reported experiencing sexual harassment, seeing the conference welcome sexist people and regularly hearing sexist or sexually abusive comments and jokes. Women