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." ∎

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

reported unwelcome, persistent advances from men at the conference. The analysis does not reveal what percentages of respondents reported these experiences, but does say that 15% of respondents were women.

Terrence Sejnowski, president of the foundation that oversees the conference, told *Nature* that the foundation's board, and others, had read the report with great interest, and thanked the authors for the analysis. "It provides us with valuable information for understanding our community," he said.

DIVERSITY MEASURES

The survey was carried out by Katherine Heller, a machine-learning researcher at Duke University in Durham, North Carolina, and Hal Daumé, a machine-learning researcher at the University of Maryland in College Park, who are the diversity and inclusion chairs at this year's event.

In December 2017, Sejnowski and the chairs of the boards of the 2017 and 2018 conferences acknowledged that several events held at or in conjunction with the 2017 conference had fallen short of the standards required to "provide an inclusive and welcoming environment for everyone". They said that they would take immediate action, including recruiting the

diversity and inclusion chairs, formalizing the process for reporting concerns and strengthening an existing code of conduct, by which all attendees and sponsors will have to abide in future.

Their statement came shortly after several female machine-learning researchers spoke out about their experiences at last year's event in Long Beach, California, and other AI conferences, including a joke about sexual assault, allegedly made by a member of a band composed of leading researchers at a party coinciding with the 2017 event.

Other measures to improve inclusion include subsidized childcare and a diversity meeting. There are also now several ways for conference-goers with concerns to notify organizers.

And on 16 November, the board abandoned the commonly used acronym, NIPS, and renamed the event NeurIPS. A March 2018 letter to the board, signed by 122 academics at Johns Hopkins University in Baltimore, Maryland, said the NIPS acronym was "prone to unwelcome puns" and revealed further goings-on at the conference, including an unofficial sister event named "TITS" and T-shirts spotted bearing the slogan "my NIPS are NP-hard".

Researchers have mixed views about

whether the board's efforts will bring meaningful change. Raia Hadsell, a machine-learning researcher at DeepMind in London who has been attending the conference for more than a decade has not witnessed a "rampant culture of discrimination, bias or harassment" at the event but has seen and experienced problematic behaviour. "I find it infuriating to be asked whether I am a recruiter, or a 'plus one', or whether I 'did the work myself' — do men ever, ever get asked questions like that?" she says.

She thinks that the machine-learning community wants to address the problems, but that their complexity makes it difficult. "I think that there will still be a problem come December in Montreal."

Elana Fertig, a computational biologist at Johns Hopkins University who signed the March letter to the board, says that altering the name is a powerful first step that has heightened awareness of the issues and shows that change is possible. But two of Fertig's students decided earlier this year not to attend the event because of the reported culture. And she worries about a backlash against the name change, noting that there were negative, sometimes threatening, comments that accompanied the debate over the change. ∎

Alzheimer's researchers seek better mice

Several teams are developing animal models that more closely mimic the disease in people.

BY SARA REARDON

Drug companies have spent billions of dollars searching for therapies to reverse or significantly slow Alzheimer's disease, to no avail. Some researchers argue that the best way to make progress is to create better animal models for research, and several teams are now developing mice that more closely simulate how the disease devastates people's brains.

The US National Institutes of Health (NIH), the UK Dementia Research Institute and the Jackson Laboratory — one of the world's biggest suppliers of laboratory mice — are among the groups trying to genetically engineer more-sophisticated rodents. Scientists are also probing the complex web of mutations that influence neurological decline in mice and people.

"We appreciate that the models we had were insufficient," says Bruce Lamb, a neuroscientist

at Indiana University in Indianapolis who directs the NIH-funded programme. "I think it's sort of at a critical juncture right now."

Alzheimer's is marked by cognitive decline and the build-up of amyloid-protein plaques in the brains of people, but the disease does not occur naturally in mice. Scientists get around this by studying mice that have been genetically modified to produce high levels of human amyloid protein. These mice develop brain plaques, but no memory problems.

"I think it's sort of at a critical juncture right now."

Many experimental drugs that have successfully removed plaques from mouse brains have not lessened the symptoms of Alzheimer's disease in people. One high-profile stumble came last month, when three companies reported that their Alzheimer's drugs — from a class called BACE inhibitors — had failed in late-stage

clinical trials. Although the drugs successfully blocked the accumulation of amyloid protein in mice, they seemed to worsen cognitive decline and brain shrinkage in people.

The drive for better mice comes as genomics studies are linking the most common form of Alzheimer's — late onset — to dozens of different genes. This diversity suggests that each case of the disease is caused by a different mix of genetic and environmental factors. "There is no single Alzheimer's disease," says Gareth Howell, a neuroscientist at the Jackson Laboratory in Bar Harbor, Maine.

Howell argues that scientists' reliance on inbred lab mice with only a few engineered mutations might have limited research. His own work suggests that, in mice, just as in people, genetic diversity plays a part in determining how neurodegeneration progresses.

When Howell's team modified two genes associated with early-onset Alzheimer's in both lab mice and their wild cousins, all of ▶