

► a research company in San Francisco, California, co-founded by entrepreneur Elon Musk, released a ‘meta-platform’ that enables AI programs to easily interact with dozens of 3D games originally designed for humans, as well as with some web browsers and smartphone apps.

All three releases provide researchers and software developers with easy ways to test programs in previously unseen situations, and for the programs to acquire new skills by teaching themselves to navigate novel situations that resemble real-world scenarios. “Environments like these have a very important role to play in the future of AI,” says Pedro Domingos, a machine-learning researcher at the University of Washington in Seattle.

#### ATARI ALGORITHM

Games have been test beds for AI for decades, but, typically, the algorithms have played them following predefined strategies. In recent years, the focus has shifted to machines that could learn from their own experience. In early 2015, DeepMind unveiled an algorithm that taught itself how to play classic Atari arcade games better than any human, by trial and error, without being told the goals of the games.

Such games are simple 2D worlds, though. ‘First-person’ 3D video games such as *Minecraft* — which visually embed the player in the environment — are a much closer approximation to the real world, and so make more sophisticated test beds.

*Minecraft* enables users to interact with virtual bricks, and use them to build structures,

in addition to navigating and interacting with predefined structures. The version now available to developers, called Malmö, allows algorithms to do the same. Hernández-Orallo, for example, is using this to explore whether the environment can be used to create benchmarks for machine intelligence. Algorithms could compete to arrange bricks into something that looks the most like a certain object, say, or to navigate a maze — testing a much wider range of skills than the Turing test, the most famous test of machine intelligence, which focuses on the ability of an AI to chat like a human.

One of the things that made *Minecraft* attractive for conversion into an AI test bed is that it already enabled players to communicate using text messages. This could help an AI to learn to collaborate with humans in the real world, says computer scientist Katja Hofmann of Microsoft Research in Cambridge, UK, who led the team that created Malmö.

#### ROBOT REHEARSAL

Virtual worlds are also particularly useful for developing AIs that are destined to eventually operate as physical robots, says Hofmann, because such environments are cheaper to customize, and faster and safer to practise in than the real world. They also allow robotics researchers to focus purely on the intelligence part of the equation — the mechanical challenges of physical robots can be a distraction.

In addition to Hernández-Orallo, Microsoft Research has collaborations with a handful of

research labs that are using Malmö projects. But Hofmann suspects that many more are using it, perhaps around 100.

DeepMind Lab similarly allows researchers to create structures such as mazes, and their algorithms can learn to collect rewards as well as to navigate. DeepMind has also been experimenting with integrating “more naturalistic elements”, such as undulating terrains and plants, into the platform, says a spokeswoman. Now that the environment is open, the company hopes that other researchers will help to make the environments more challenging for the algorithms. “By open-sourcing it, we are allowing the wider research community to get involved in shaping this,” she says.

OpenAI’s meta-platform, Universe, takes things even further. By providing multiple, radically different environments for the same AI to sample, it could help to attack one of the hardest problems in the field: creating algorithms that can use previous experience when faced with new situations. For instance, deep neural networks, which mimic the layers of brain cells in the visual cortex, can quite quickly learn to navigate a 3D maze, but cannot transfer the knowledge to navigate another maze. “If you change the colour of the maze, the system is completely lost,” says Hernández-Orallo. “State-of-the-art technology fails dramatically.”

Microsoft is now working to make Malmö available through Universe. “Having a community platform will accelerate everyone,” says Greg Brockman, co-founder and chief technology officer of OpenAI. ■

#### POLITICS

# No fond farewell for Italy’s premier

*Scientists feel let down by ex-Prime Minister Matteo Renzi.*

BY ALISON ABBOTT

Italian politics is in turmoil after the resignation of Prime Minister Matteo Renzi — but researchers say that they are not particularly sad to see him go.

In his almost three years in charge, Renzi promised improvements for universities and science but failed to raise the status of research in the country, according to scientists who complain that he also directly interfered in academic affairs.

“Renzi became prime minister at a time of serious economic and social crisis, and he

injected a sense of energy and optimism into the university and research sector,” says biologist Cesare Montecucco of the University of Padua. “Our expectations were raised, but they were mostly disappointed.”

Renzi resigned on 7 December, three days after constitutional reforms that he proposed were defeated in a referendum. He stayed on to push through a 2017 budget that sees no significant increase for Italy’s chronically underfunded university and research system. (Exact figures for research spending have not yet been released.) Italy’s research and university funding per head is among the lowest in Europe

— although the country does produce a greater share of highly cited research papers than the European Union average. Little has changed on that score during Renzi’s tenure, say Montecucco and other scientists.

Renzi has not delivered what they have long campaigned for: less bureaucracy for research institutions and a new research-grants agency along the lines of the US National Science Foundation.

#### FUNDING FALLACIES

Most controversial has been Renzi’s November 2015 decree creating a €1.5-billion (US\$1.7-billion) centre for genomics in Milan. Known as the Human Technopole, it will focus particularly on personalized medicine and nutrition. The country’s 2017 budget foresees annual funding of well over €100 million, beginning in 2018.

Although some are grateful for the research funding, many scientists have complained that this major investment in a single new project is inappropriate when most other public research institutes are starving for cash. They also strongly objected to the fact that it was planned by Renzi with a few

## METRICS

chosen scientists, behind closed doors.

In September 2016, Renzi floated the idea of creating 500 elite professorships known as Natta chairs (after Italian chemist and Nobel laureate Giulio Natta), to be awarded mainly to Italians working abroad. They would be selected by 25 evaluation panels whose chairs the prime minister would nominate. Thousands of academics signed an open letter in October complaining that Renzi had designed the programme without discussing it with universities. The letter also protested against the involvement of politics in the selection.

Regulations for the Natta selection procedure have not yet been published, and so scientists hope that the next govern-

**“This means that weaker universities in the south will lose even more money.”**

ment will ensure that the process remains inside the academic community. “Nomination of panel chairs by the prime minister is just not acceptable,” says physicist Giorgio Parisi of the University of Rome La Sapienza, a prominent critic of the process. “It is a political choice to do the selection independently of Italian universities, but then you could turn to external academic organizations, like Europe’s national academies.”

## BUDGETARY BLUES

Parisi is also unhappy with aspects of the 2017 universities budget. In particular, €271 million will now be reallocated to the university departments that are judged by the national evaluation agency ANVUR to have the best research performance. Parisi thinks that rewards for high performers should come from new money, rather than being transferred from a general university budget that is already stretched thin. “This government reallocation means that weaker universities in the south will lose even more money, and this would be a social disaster,” he says.

An interim government will hold down the fort until new elections are held, which could take place next year. Uncertainty is set to continue. Populist and protest parties, particularly the Five Star Movement led by comedian Beppe Grillo, are likely to make substantial gains in the next election.

These parties do not have strong scientific agendas. Italian senator-for-life Elena Cattaneo, who is also a neuroscientist at the University of Milan, is taking a wait-and-see perspective. “One or two populists in the current parliament have shown themselves to be more open to discussion on scientific topics than members of mainstream parties,” she says. ■

# Impact factor gets heavyweight rival

*CiteScore uses larger database and gets different results.*

BY RICHARD VAN NOORDEN

One of science’s most contentious metrics has a flashy new rival. On 8 December, publishing giant Elsevier launched the CiteScore index to assess the quality of academic journals.

Although the index ranks journals with a formula that largely mimics the influential Journal Impact Factor (JIF), it covers twice as many journals — 22,000 to the JIF’s 11,000 — and its formula includes tweaks that produce some notably different results. These include lower scores for some high-JIF journals (see ‘A new measure of journal impact?’).

If CiteScore becomes popular, these quirks could change the behaviour of journals hoping to maximize their score, say analysts. But CiteScore’s debut comes at a challenging time for such metrics. It’s not obvious that there is an appetite for a competitor to the JIF, and scientists note that no matter what differences CiteScore provides, it will have to survive the same criticisms that are lobbed at its rival — most notably that the JIF is so commonly promoted by publishers as a yardstick for ‘quality’ that researchers are judged by the impact factor of the journal in which their work appears, rather than by what they actually write.

“In my view, journal metrics should always be accompanied by health warnings that are at least as prominent as the ones you see

on cigarette packets,” says Stephen Curry, a structural biologist at Imperial College London. “Such metrics are at the root of many of the current evils in research assessment.”

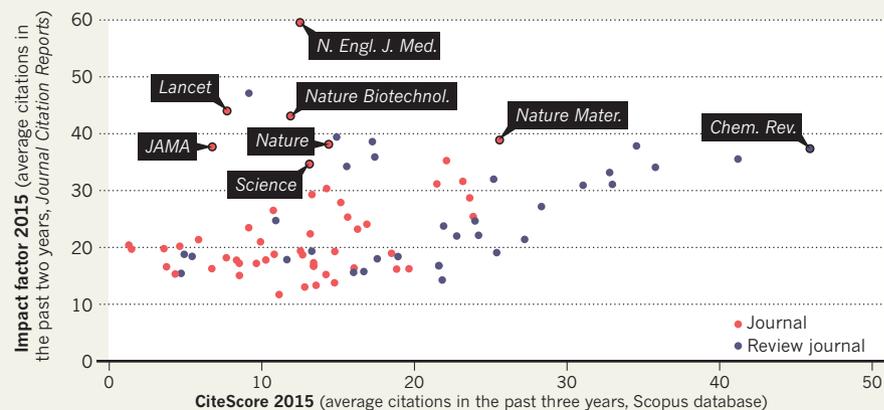
Amsterdam-based Elsevier has for many years provided a suite of analytical indicators, including journal metrics that have never become as popular as the JIF. It says that it has launched CiteScore owing to “overwhelming demand” from authors and editors.

The publisher is uniquely placed to challenge the JIF’s hegemony. It owns the Scopus database, a record of article abstracts and their reference lists. Aside from Web of Science, on which the JIF is based, it is the world’s only reasonably comprehensive and carefully curated citation database. But Scopus is bigger, enabling scientists, librarians and funders to check the popularity of many more journals. Furthermore, unlike the JIF, which is available only to subscribers, CiteScore figures will be free online for anyone to view and analyse, although full details of the documents included in the calculations are visible only to subscribers.

When it comes to their underlying formulae, CiteScore and JIF are near-doppelgängers. To score any journal in any given year, both tot up the citations received to documents that were published in previous years, and divide that by the total number of documents. The most popular version of the JIF looks at research articles published in the previous ▶

## A NEW MEASURE OF JOURNAL IMPACT?

Journals that have a high impact factor, a measure of the average number of citations that their articles receive, don’t necessarily score so well on a new indicator, CiteScore. The latest metric includes documents such as editorials, letters and news items, which attract fewer scholarly citations.



The highest-scoring journal on both measures is CA: A Cancer Journal for Clinicians with an impact factor of 138 and CiteScore 66.

SOURCE: JCR/CITESCORE METRICS