



Scientific research increasingly relies on custom-built software.

GETTY

# EX-GOOGLE CHIEF'S VENTURE TARGETS NEGLECTED SOFTWARE

The US\$40-million Virtual Institute of Scientific Software will fund the maintenance of researcher-written code. **By David Matthews**

**S**ee whether this sounds familiar: you build a piece of software to solve a research question. But when you move on to the next project, there's no one to maintain it. As it ages, it becomes obsolete, and the next academic to tackle a similar problem finds themselves having to reinvent the wheel.

That's what happened with WaterDetect, a tool for identifying water in satellite images. Mauricio Cordeiro, a geospatial data scientist and programmer at the Laboratoire Géosciences Environnement Toulouse in France, revived the moribund software at the

start of his doctoral project, spending some four to six months getting it to work. Now in the final year of his PhD, Cordeiro still devotes considerable energy to the program. But after he graduates, he says he won't have time to address issues that arise, and there is no one at his laboratory who can take on the project for him. "The problem continues," he says.

Now, a funding initiative hopes to help ease that burden.

In January, Schmidt Futures, a science and technology-focused philanthropic organization founded by former Google chief executive Eric Schmidt and his wife Wendy, launched

the Virtual Institute for Scientific Software (VISS), a network of centres across four universities in the United States and the United Kingdom. Each institution will hire around five or six engineers, says Stuart Feldman, Schmidt Futures' chief scientist, with funding typically running for five years and being reviewed annually. Overall, Schmidt Futures is putting US\$40 million into the project, making it among the largest philanthropic investments in this area.

The aim is to overcome a culture of relative neglect in academia for open-source scientific software, Feldman says, adding that support

for software engineering is “a line item, just like fuel” at organizations such as NASA. “It’s only in the university research lab environment where this is ancillary,” he says.

Feldman clarifies that VISS engineers will primarily be supporting existing Schmidt Futures researchers who “desperately need good software”, rather than taking on new awardees. But the hope is that they will be able to lend their expertise more broadly.

The institute will initially fund projects at the University of Cambridge, UK; the University of Washington in Seattle; Johns Hopkins University in Baltimore, Maryland; and the Georgia Institute of Technology (Georgia Tech) in Atlanta. It’s still early days, but several software projects are already lined up for support, Feldman says. One involves a data-management system to provide researchers and policymakers with access to confidential government data; another will analyse data from a proposed radio telescope.

At Cambridge, VISS resources will help to rewrite the software behind the Scale Aware Sea Ice Project to better model the impact of climate change on sea ice. Georgia Tech is weighing several projects, including the Virtual Ecological Research Assistant, which simulates ecological changes. Meanwhile, the University of Washington has been “flooded” with requests from researchers there who say they need engineering help, says Sarah Stone, a data scientist at the university who is helping to coordinate its VISS centre. Washington is looking to hire a lead, deputy and five primary software engineers. They will range from graduates fresh out of their undergraduate programmes to multi-year industry veterans.

## Closing the gap

Those setting up VISS centres say Schmidt Futures’ steady, relatively long-term funding will help them to overcome a range of problems endemic to academic software. Research grants rarely provide for software development, and when they do, the positions they fund are seldom full-time and long-term. “If you’ve got all of this fractional effort, it’s really hard to hire people and provide them with a real career path,” says Andrew Connolly, an astronomer who is also helping to set up the Washington centre.

What’s more, software engineers tend to be scattered and isolated across a university. “Peer development and peer community is really important to those types of positions,” says Stone. “And that would be extraordinarily rare in academia.” To counter this, VISS centres hope to create cohesive, stable teams that can learn from one another.

Still, there’s a limit to how far Schmidt’s investment can go. Hiring good engineers in places like Seattle and Cambridge is expensive, because universities are competing against the likes of Amazon, Microsoft, Google and Meta.

“It’s tricky,” says Chris Edsall, head of research software engineering at the University of Cambridge’s Research Computing Services. “We can’t unfortunately compete on salary.”

Instead, VISS is banking on the lure of the research itself. “One of the things that we’re hearing from software engineers is that it’s not just about the money, it’s also about the impact that they can have,” says Connolly. He hopes the VISS centres will create a “semi-permeable membrane” between academia and industry, tempting engineers to return to campus.

Yet even those establishing VISS centres admit that the needs of the open-source community far outweigh the investment. The scheme, although welcome, is a “drop in the bucket”, says Irfan Essa, a senior associate dean at Georgia Tech’s College of Computing who is helping to set up the VISS centre there. “It’s a pilot.”

Dario Taraborelli, who helps to coordinate another privately funded scientific-software project at the Chan Zuckerberg Initiative

## “They don’t survive this ‘Valley of Death’. The funding stops when they develop the software prototype.”

(CZI) in California, says that such initiatives fill a key gap in the scientific-software ecosystem, because funding agencies too often fail to prioritize crucial software infrastructure. Although there are now “substantial” grants dedicated to creating software, he says, there’s precious little funding available to maintain what is built.

Computer scientist Alexander Szalay, who is helping to set up a VISS centre at Johns Hopkins, agrees, noting that very few programs get to a point where enough researchers use and update them to remain useful. “They don’t survive this ‘Valley of Death,’” he says. “The funding stops when they actually develop the software prototype.”

To counter that, CZI has since 2019 distributed \$27.8 million as part of its Essential Open Source Software for Science initiative, the fifth round of which was announced in March. Among its beneficiaries is NumPy, a numerical and scientific computing package for Python. CZI funds have helped the project to expand its pool of contributors by 60%, add its first female maintainer and attract code from contributors in Africa, Asia and South America, says Taraborelli.

Private companies can also play a part by developing commercial software packages, says Elliot Hershberg, a genome scientist at Stanford University in California. He says that some even offer “Robin Hood-esque” models, charging industry full price while keeping the software free for academics. But private sector

involvement risks a “lack of transparency” as well as a “huge subscription cost” for academics, he warns.

## Culture shift

As demand for scientific code increases, the hiring culture around software in academia has begun to change, says Alex Orso, another computing associate dean at Georgia Tech who is involved in setting up the VISS centre. In a reflection of the growing importance of software in academia, scientists are perhaps just as likely to boast of their GitHub prowess as their publication rate. In a recent interview round for a position outside the VISS, Orso notes, candidates “all were showing the number of ‘forks’ and ‘stars’ on their GitHub projects”, referring to metrics that reflect the popularity of their work. “That is something I wasn’t definitely seeing a couple of years ago.”

The position of research software engineer (RSE) has likewise become increasingly valued. RSEs are specialists who work on software specifically to advance academic research; they’re like at-large developers for a department or school. Since it was founded in 2018, the US Research Software Engineer Association has expanded its membership to more than 1,000. In the United Kingdom, at least 31 universities support RSE teams. “I really do think this a growing area,” says Daniel Katz, an associate editor-in-chief at the *Journal of Open Source Software*.

Development is supported by an influx of new federal funding opportunities. In 2021, the US National Science Foundation launched a new type of call, backed by \$4 million, to help software projects transition to a sustainable basis. Since 2020, the US National Institutes of Health has awarded \$6 million to \$7 million per year to ensure that software created by existing awardees is robust, sustainable and scalable.

The German Research Foundation has also launched funding calls focusing on software sustainability and quality. And in the Netherlands, there has been a rise in hiring as attention has turned to the need for quality research software, says Maaïke de Jong, a scientific community manager at the Netherlands eScience Center in Amsterdam, which supports research software through funding projects and training. Staff numbers have nearly doubled in the past three years, she estimates. “There’s job interviews all the time.”

But as ever, the need outweighs the funding. As part of its investment, the Dutch government is providing funds to create digital competency centres at the country’s universities, many of which are hiring RSEs, says de Jong. But awards were capped at €250,000 (\$263,000) per research organization, just about enough to hire one or two people for a couple of years. “There is still a long way to go,” she concludes.

David Matthews is a freelance writer in Berlin.