Q&A Alice Allen

The code librarian

By day, Alice Allen runs software and IT training programmes for the Board of Governors of the Federal Reserve, the US central bank, in Washington DC. But in her spare time, she edits the world’s largest registry for software in astrophysics and astronomy research — the Astrophysics Source Code Library (ASCL; ascl.net).

How did you come to edit the ASCL?
I’ve always been interested in astronomy. So in 2010, I started doing volunteer work for the website Astronomy Picture of the Day. I told one of the site’s creators — Robert Nemiroff, an astrophysicist at Michigan Technological University in Houghton, which hosts the ASCL — that I had time for another project. We talked about an effort that he and John Wallin (now at Middle Tennessee State University in Murfreesboro) had started in 1999, to create a repository of astrophysics source codes — the old ASCL. The resource had gathered about 40 source codes, but lay fallow at the time for want of an editor. I took it over and have been working on it ever since.

I work on the ASCL in my spare time and I take vacations to speak about it at conferences. It’s an all-volunteer organization: we have two developers and an associate editor, Kimberly DuPrie, a programmer at the Space Telescope Science Institute in Baltimore, Maryland. The advisory committee is made up of astrophysicists who do this because they think it’s a good idea. There’s a lot of passion around this project.

How do entries get added to the library?
For the majority of entries, Kim and I look through research papers specifically to find codes to register, and we e-mail the authors; we know that coders won’t necessarily think to register their codes with the ASCL. But since a site redesign in 2014, about 40% of the listings have come from scientists submitting their own entries, which we verify. We don’t look at the quality of the code, but it has to meet our criteria, such as being used in research and immediately available to download.

If a link goes bad, we track down the code’s new home. Journals don’t have to worry about references to websites going stale: their links will stay good because they point to the ASCL.

How is it funded?
It’s mostly unfunded. We have in the past received a few thousand dollars to cover things like conference presentations and poster production costs, from organizations such as the American Astronomical Society and the Heidelberg Institute for Theoretical Studies in Germany. Sometimes my travel to conferences is funded, but often it is not. Over the long term, the goal is to make the ASCL financially stable, although a large part of it will always be volunteer-based.

Do other science disciplines have equivalent code libraries?
There are lots of code registries and repositories online, but they’re usually not science-specific. We index codes that are useful specifically in astronomy and astrophysics. There isn’t one site for all science codes, and maybe there should be.

I’ve had several people approach me about creating a similar resource for physics. So last December, we started offering clones of the ASCL infrastructure (which runs using open-source software) to any discipline that wants to build a code registry of its own. A sample site is at sci.codes.net, and Michigan Technological University is willing to host other science-code registries for three years if people would like. Nobody has taken us up on this offer yet, but they are welcome to try it out.

BY JEFFREY M. PERKEL
This interview has been edited for length and clarity.

— publishers, usually at no cost.

Coles himself covered the costs of developing the software platform for the journal, amounting to a few thousand pounds, he says. (Discrete Analysis licenses different software and is helped by a grant from the University of Cambridge.)

GitHub is covering the costs of hosting the platform, so the only remaining expense is editors’ and reviewers’ time, which they give voluntarily, says Coles. If the experiment proves successful and the volume of papers balloons, the journal may eventually have to charge authors a handling fee of a few tens of pounds, he adds. (The journal also relies on the continued existence of arXiv, whose running costs amount to less than US$10 per paper).

The journal does not have the resources to offer services provided by conventional journals, such as heavy editing of papers. Instead, poorly written articles will be rejected and the authors referred to a list of professional copy-editing services, Coles says.

GAINING TRACTION

Gowers welcomes the new journal; the arXiv-overlay model is much more likely to succeed, he says, if many examples of it can be seen to be working. The journal has amassed an editorial board with high-profile physicists including Pedro Ferreira, a theorist at the University of Oxford, UK, and Andrew Jaffe, a cosmologist at Imperial College London.

But astrophysicists will not necessarily jump to publish in Coles’s journal. Ewine van Dishoeck, an astrophysicist at the Leiden Observatory in the Netherlands, says that she, for one, is unlikely to submit her work there. “We have a small number of well established and high-quality journals in astronomy that everyone respects,” she says.

Papers in astrophysics are effectively open already, van Dishoeck points out, because anyone can view preprint manuscripts immediately on arXiv, whereas journals in the field make final accepted versions open after a delay — typically 12 months after publication. An issue for journals in the field make final accepted versions open after a delay — typically 12 months after publication. An issue for journals hoping to achieve traction among researchers are ensuring speed and editorial fairness, adds Andrew King, a cosmologist at the University of Leicester, UK. “Reliability — and particularly fairness — are very hard to guarantee,” he says, pointing out that the backing of long-lived organizations with a stake in the future of a field, such as learned societies, is often crucial to a journal’s success.

TOOLBOX

How are code libraries used?
The code librarian at the University of Leicester, UK, Federico Disalgo, says the online registry /astrophysics.hedmonds.com has been cited more than 500 times since 2012, and the #code directory that he hosts on GitHub has more than 100,000 hits in 2014. It provides a way to tie together the pieces of results. The ASCL holds nearly 1,200 records of source codes that have been used in research in peer-reviewed publications. Each entry is citable with a unique ID and points to a website where the code can be downloaded; entries include a description of the code, its authors and some of the research it appears in. We also house some source codes. The site got more than 100,000 hits in 2014. It provides a way for journals to point to a software record — it has been cited more than 500 times since 2012, according to NASA’s Astrophysics Data System Abstracts Service.

Do other science disciplines have equivalent code libraries?
There are lots of code registries and repositories online, but they’re usually not science-specific. We index codes that are useful specifically in astronomy and astrophysics. There isn’t one site for all science codes, and maybe there should be.

I’ve had several people approach me about creating a similar resource for physics. So last December, we started offering clones of the ASCL infrastructure (which runs using open-source software) to any discipline that wants to build a code registry of its own. A sample site is at sci.codes.net, and Michigan Technological University is willing to host other science-code registries for three years if people would like. Nobody has taken us up on this offer yet, but they are welcome to try it out.

BY JEFFREY M. PERKEL
This interview has been edited for length and clarity.